



1/118

FIG. 1A

The Small Island: 10848bp (SEQ ID NO: 1)

TTTTCGAAGGGGCGAGTTGCACAGGCCTAGGAAGGGGTGTTGCACCTTTTACTTGTCCAGCAGCCGTCTGGGCCGAACATTAGG
CCCCGATTGACGTGATGTCAATATGGAATAGCCCGCTATTAGTGGGCTTTTCTGTTTTTCAGGCCCTGAAACCAATTCTGCA
GGCGTAGCAATGTGCCCTCGACGTAATCAGCGATTACAGAGGGTACTGCGGTTTCGGCCTTGAAATTGCGGGGCTCTGCCAGCA
TGCCGGCGCTGGAATCCAGGCGCTCTCGGTAATCCAGATTTTCTACAGGCCTAGCGAACTGCTGGGAGGCGCAGCTACCTT
TCGTCCGGTAGCCGCCCCCTCAGTATCGGTATCGGTAGGCGAGACTGCTTCGGGCGGTAGATCCCCATCTCCAGATCCGCTTG
TTGCGGTGGAAGCCGCGGAGGCGTGGCGGCACGGTTCGATAGGTGTTGAGCCATCAGAGTGAACAGCAGCGTTTGTCTCGAAGG
TGTAGGTGGCTGTGGTAGCGCTTCTTCCACGCGGTTGCATCTCCCTGAATGCCTGTCCCTGGTGTGTGCTCCCTTGGTTCTG
CAGGTTCTGTCTAGCCTCTGTGGCGGCGCTTACCCATTCGAGCGAGAGGTGTCGCCGATGGCTTCGGTTTGGCAGAGGCTT
GCGGGGAGCCGCTGTAATCCACCTTCCGACGTTTCCAGCGTCTTGGGCCAGGCTTCTGGCATGCAGCAGGCAACGATGT
GGTCGTAGTCGGCCTCGATGTAGCTCATGGTGGTCTGGATATTCGAGTGGTTGAGCAGGCACTTCGTGAGGTGGATGTTCCGCT
CGGGTGCTTTCATCAAGTCGGTGGCCAGGCTGTGCCGGAACCGGTGCGGAGTCATCCGACCCCAACCTTCTCGGTCAAATTC
GGTACATGGCTTCGACCTGGTCGGAGTTCATCACCTTGCTCTGTAGTGCGGTGAGAACCGGTTGACGTTGAACAACTGGTCGT
CATCCGCGTAATCGGCTCTATCGGCTCTGTCAGGAGCCGCGTTCGAGTGGCAACAGCCCTCGGTTATGGCACTCGAACT
CTTTATGAGTTTCTCTGTCTCGCGCGGATGAGGATCAGTTGATTTTCCAGTCGATGTCCGCTTGGCGGATGCAACAGCG
CATTCAACCGGATGCGCGTGAAGTAGAAGACCTCAACCGTGCAAGCCAGAACCCAGGCGGGAGTGATCGGTGCGCGATCGCCAG
TGCAGCGCTCGCGCGGACCTGCATGTTGAGCCAATTGCGGGCGCGCAGGATGGCTTCGGCTGCGACGGTTTTGCTTGTCTGCC
TGGGGGAGCCGCTGTAATCCACCTTCCGACGTTGATTTGGGATGTGTACCAGCTCTGTCTGAGCAGGTAGGCTTTGAGCTTTTCTGCTT
TTCGAGATGATTGCAATACGTGTTCCAGCTCCGCTTCGACAGGCCTTGTTCAGGACCTTGCGCCGCCATCCAGCACAGACC
TGTGGTCCACCTCCTGTACGGTTGCCGTAGGGCCGAAGTGCTTGAACAGCGCTTGGTTCGGGCGCGGTAGATCTTCGCGCTGG
CTTCTCGGAGATCGTGCAGGAGATGTAATCTCCTCGGTGAGCTGCTGCGGCGCTCATTCCACACCTCCTTCGGCATCGGTGATGAC
CGTGAGGCTTTGGGTTGTCAGAGGCTGCACAGGGAACGACCAATTGGGATCTCTGGAGCAGGTAGGCTTTGAGCTTTTCTGCTT
GCGAGGACAGAAACCTTGATGGTCCAGATAATCAGGTTCTTACTGGTCTTTCGGTGAAGACCTGTTTCTCGAACGCGCGCTG
CACCAGCTTCCAGCGCGGCTCTCCTTGGCCTGGGCGAGCTTTTCAACCTCCGGATGCTCTGGACATAACGCTTGAAGATTCC
TGGCGTGACCCAGCATGGCGGTCCGCTCGACGGTATGCACCAAGCCCTTGGTGTCTGTGATGAACAGGCGACGGGTGCGGATACC
AGATTTTCATCCAGCCAATGAATCCCTGTCCAGATCAGTCTTCTAGGCATAAATACATCTTCAATTGCTTCAGGAGCGCA
ATCTGTAGTGGTCAACCCATCCCGGACAGTGGGTGAGTTTTTCTTCGGCCACCGCAGGTGGTAACCCGCTCGTTGAATTGATGC
GGCTGATCCAGTTGTAGCGGTGCATCAAGTAATGACTGATGTCCGCTTGGGCTCCTGCGCCGTACGGTAACCCGTTGACGGG
ACCCACTCCGACTTCAGACTGCGGAACAGGCGCTCCATCGCGGAGTTATCCAGCAATTCCCCCGACGGCTCATGCTCTGCTGC
ATCCGATAGCGCCAGAGCCGTTCGCGAAACAGGCGCTGGCGTACGCTGCGCTGGTCTGAATGGAACAGCACCTGCTGTGGC
CTGCCGCGCTGTTCTGATGGCCATGTCCAGGGCTTGTATCAGGTTTCGGCATCCGGCTTGGCCGAGAACCGCCAGCGCATCAGC
AGATCCAGCACCGCGGCCAGGTAGTGCAACGGCCTTGCGCCAGACGTAGGTGATGTCCGCAACCAACACCTGTATGGGATGC
TCGGTCCGCAATTCCGGGTTAGCCGATTCCGGGATATCCGGCGCTCAACCGTGGCCTGTTTGTAGGCGTGCAGCCCGGTTGC
TTGCTGACCTGCGCAATGAATCCCTGTCCAGATCAGTCTTGTGCTGCTCAGGCGTCAACGCTTGTCTTCGGGCTACACCTGTGGC
GTCCGCGCTCGAGCTGCTTCAACCAACGGCGCAAGGCCAATCCACCACCCCGAGCAACGGCAGGCGTGCATATGGCTGTAGC
CTTGGTCCAACACAGGGCGCGGCTCTCGTTTGAATCGGCGGAAACGTAACGTCTGCTTGTCTCATCAGACACCTCTTTC
ACGGCGAGGATTTCTCGCCTAAATCGGTGTCCGGGATCAGTAGCACTACACTGCGCAACAGTATTTTGTGCTAAGAGGC
AGGAAGGCTGCGGAGATTTCTACTAGTGAATTAGATGTTGTGCTTCTGAAACAGAACTGAAGTGAAGTTGGGGTAGGG
TTTTTCTAGGTGAAGGTAACCTCTTGAGAATTACAAAGGTGTGACGCAATGTTGTATTTTCTTGAGATGAAGATGGGTGG
TTGGGTCCGATATAGGTACTTCTCTATTTTCTTAAATTGCTCTCATCTATGGGTGTGTGGTGGTGGAGGTGGATCGGATGA
GATTGGGCAGCACTGCTTGTGAGAGAGAGCAAAAGCTTTCGGAGTTAATGATAATGAAGAGGGGAGTGTGAGGTTGAATCGGCT
GAATCGCATCCGAATTGAAGGTCTGTTTGAATCAGAGAAGTGAATAAGAAAGCCGCCAATGAGTGGGTATTCAGTGAGT
TAATAATTTTGTAGCTGAGAGGTCTTTTGAATAAGAAATCACTTGTATGTGCGGCTGTGCTTTGGTGGCTAGCAATTT
CGCGTGTGCTGATGAGGGCTCAAATGATGGAAGTGAGATATGTGGGCGCAGGTTGAGTTGAAATAACAAGTCTGGGGGAAGT
CTCAAAGGGTGTGGATGTTGAAGATGTTGTAGTTTGTTCGATTCTTCCAAGTAATATGAAGTCAAGTCAAAGAGCGCCTACACT
CCCTCCTCTGCAAGAGATGATCATTTCGGCAATGCTTCAACCGAACGGTCACTGTTTTCGACGCGGAGATAGGAAATTTAC
AACATCTTGGCGGGCAATCTTATGTCTCCACGTTATGCCAATTTCTATCCAGACGGTGTAGCAGGGGAACATCAGATCTACG
ATGTGTTGGTTACAATACACCGGGAATTCATCTCAAGGGTGAATGTGTCTATGGGACGCGCCGACGACATTCAATTGGGTGT
TGAGCCATATGGCGGATCTGTTGTTGTAACCTACAGTGTCACTGCATTCAAAACAACGATTCCAGTGAATAGCTACAGTTA
TCGTGATGGGCGGGCAGTGTATGGCGAGGTCCAGAAATGTGTGAGGAATAAATGTGGTTTGAACCTAAGGATAGGCTTAGT
AGGATTTTATTTCAAGGTGTTGAGCTTGTGAATGTGTGACAGGCACTCTATAGAGTGCCTGATAGCTGGAAGTCTGTTTGG
CTGAGGGCTAGTCTTTCAGTTCGATATGGCACGTTGATGAAGGTGGTAAAGTCTGTCTGTGTGGTGGCAGGCGAGTGG
CGGCTTCAGACGCTAATGTGGAGTTCAAGGTAGGTTCCGACGCGGAGGATCTCTTGGGTGGAGTCAATTGCTGAGCCATGC
AGTTATAAGACAGGCTACGTGCTCCATTGACTTGGACATCGATACGAAATTCAGATTGTTCCGAGTGTGACTGTCGGGCG
GGGAGGCGAGCAGCTGCTGTGCCAAACCACTATAGCAACAAGAAATAATTCGAAGAATTTAATAAGCAATTCGAATCTCCT
TAACCTGAGGGTTGGATGAGACTCAATAGTAACCTCAAGTTATGAGGAAAGATGAACCGCTATCGTTTGTAGTCTTGGTTA
GGTCAGGCACACGCTAGTGAAGTGAATAAGAGTAAGCCGATAGCCAGGGGCGTGGTGGCCCATTTGGGGAAGTGAAGAG
GTGCGGCTGTGACCGGTTGAGGAAGGAATAGACAGCAGCAGGATCGCTTGGATGGGCGAGGAAACCATCAGCCGCTC
CGTATCCAGAACTGCTGCTTGCAGCGTTAGGGGTGACTCCCTCATGCCCTAGAGGGCATAATCCCTGTATGGAGATGGCTAACT
GATATTAAGGAGTGGTGAAGTCATGGAACGCTTGTCTGAGAGCAATTACATCAATGCCCGGCGCGGATGAGTTGAGGCTTA
GCCTCACCAGCTCCGCGCGCAAGAGAAATGGTAAGATTTGTGGATGGGAGGAGGTGAGGTTCTGCCAGGTGAAGTGCAGGGCA

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND
PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September 12, 2003 Serial No.: Not Yet Assigned

Page 1 of 118

Customer No.: 21559

2/118
FIG. 1B

TCCTGGAGGCCCAAGAGGGATGTTGGAATCCTCGCGCACTTCTAGCCAAGAGTCTCGTGGCGCGACGCTAGAGATCCTTGG
 CTAGCCAGTCTATCGCCGACAGCAACGTGGTCTCGTGGCCATGCCTTCAAGCTTACGCTTCGTCTGTCTGGTGTCCCGGA
 TGACGAGGGTCGGGGTCGCTGTAATGCCCTTCGACTGTGCGGTGTCGATGTTCTGTAAGCGTCCGGGGAAGGCACATTGCGATGC
 GCTAAGCTGTAATCCTTTAGGATTGAGGTCTCAGCATGGAATGCCACGTTCCGTCCCGCCACGAGCAGAGATGCAGCAGCGATAA
 GCTGCGTAGTTTATAGCCGCCCTGCGTGAGTCAAATTCACAGGACTATCCGCTGATGTGATCGCTCAGGTTGAGCAGAGCTTTT
 CTCCTGAAGCCATCACACACAGCTTACGAAGCGTAGGGTCTTCGTAGCCTTATTGGGCGAAAAACATTATTGGCACTGCCGGTC
 TCGACGGTGACGTGTCAGAAAGTGTTCGTTGACCCAGCTCACCAGAAAGCGGTATCGGGCGGCATTGATGGATGTCTATTC
 ATACAACTGCTGCCAGCGCGGGAGTTGGAGCTGTACGTGTCCATCGTTCGATTACAGCTGAAAGGTTTATACCGCATTGGGTT
 ATCAGAAAATCCGCGACGAGTTTCATGGGGCGGAGCGCACACTCGTTATGGAGAAGCGGCTGAGGATTATTCAGACTATTTA
 CCCACTTATGCGGCAACAAATCAGTGATCTCACTGGCCCGCTGCGTGGGCGAGCCGCTCAGCATCTTTGAGATAGGCATACG
 GATCATGCCATTATACGCGCCGACTAGATCAGGCTCATGATTGAGCCGCGCCGTTTACCACTGCGCAGCGACCCGGCAACA
 GCCAGTTCGAACGCCCGAGCGCCCATGGCCGTATTGGTTCTCGACCTGATTGTTATCGATGGGCAAGCCCCATCGTCCAGGT
 AGCGCGTCAGCGCTTCCAGCGTTTTCAGGCTGTAATCAGAGGCTTGGCCGTGGCTGATCCATTGGGCAACAGGTCTCGTGAG
 CCAACATCCAGTCATGAGTTTGTGAGGATCGGCACCGCAATTCTGTGCTATTGCGCAGCGCTTTCATCATCTCATGTCCC
 GCGCTGACGTTGCACTCGTACAAGCCGCTGATCGAGTGCAGGGCTGTTAGCCAGCTGACTTTTGTTCGCCACGTGCAAAAT
 CAAAGAACTTGGCCCGGGCGTGGGCCATGCAGCCGATTTCAGTGATGCCTTGTTCGAAACCGGCTTTGTAGCCAGCGAAGTCGT
 CGCAGCTCGGCAACTTGGCCACCATCAGTGGGCCAAACAACTGCGGCTGGGATACCTTGTTCGATGACCTGCCCCGAGT
 CTTTGGAGCCCTGAAAAAGCGGCTCGTGCTGTACGCGCAGACATGCGCCGGTGGGTTTTCTTCTGCGCGGTACAGCGGCA
 CCGGTGTTTTCATCAGCGTGGATCACGCCCTGGTTTCAGCACGGCTTCACGAGTGCATCGACAGTGGCTGAAGCCGCA CGCCGG
 TTTGTCCGACCCACTGCGCCAGGGTCGAGCGAGCAATTGCCAGCCCGCGCGGCCAAAGATTTTCTCCTGCGCGGTACAGCGGCA
 AGTGATCGGCAACTTGGCCACCATCAGTGGGCCAAACAACTGCGGCTGGGATACCTTGTTCGATGACCTGCCCCGAGT
 TCAGTACCAAGCTGAACTTAGTAGAATCCGTTTTTCAAGCAGGAGACGGCAGTGCAGAAAGCGTTCTTCTCGACTCCCTCAAGCG
 GGTGAGGCTAGTGATCCGACCGAGCTGCAGCCCAACTGGTGGAACTGGCCAGGAGTATCGACGCTGTAGCTAACGCCGCT
 GCACATCCTGCTGTGGCGTGCAGGTCAACCAAGCGGATTTACCGGCTGTACCGAGCCCGCGGCTTGATGGTGAAGCGCGCGA
 GGGTCGCGCTGGCGCGCAATGCCTGAGCCTGTGAGCGCACCGAACAGGTCTTGTGATGGATTTCGTCTTCGACGCGCTCA
 GCACCTGGGCGACGAGTCAAAATGCTGACGCTGGTGGTCACTTCAACAAAGGTGTCGGTGCACATCTTGGCGAGCTTGGCTGCA
 GCGGTTTTCTGTGTACGCGGCGCTGGACGAGATGGCGGTTTTCTGGCTACCCGAGGCGATCCGCA CCGACAGGGCCCCG
 AGTTACCGGCAAGGCGCTTGATCAGTGGGCTGTGAGCGTGACATCAAGTTGAAGCTGATTACGCTGGCCAGGCCACGCGAGA
 GCGCTTTCATCGAGTCATTCAACCGCAAGTTCCGGGGCGAATGCTCAATGAGCACTGCTCGCTGGTTCGAAGCCGAAATCCGTA
 TCGCGCTTGGCGGGATTACCAACGAGACCGACACGCGCAACGCGCAATTTGGCAATCTCTCCCGGCGAGCTTGGCTGCAAGTGGC
 GAACCAACAGCAGCAGCTGAAGCGGGAAGTTGATATCAACCCCATAGCCTACTAAGTGGCAGCGGTAATAAACTGGGGG
 CAGGCCAGTCCGAAGTGAATAAAAAACGTCGCAAACTGGCAAAACCTTATGGCCGGTTTTCTTAATATTTGCTCCGAACCTCG
 GTATTTCAAGGAGCGAATGCATATCCAATCGTTGGGGCTACTGCTCCTCGTGAATCAGGAGCTGTGCAAAACCCGTCG
 CAGCTGGGCGACAGTCCGCTCGCTGCGGCTTTCAGGCAAGGTCTCGGGGTTCGCTAAAGAGCAGCTTGGCTGCAAGTGA
 CTTTCCGGGAAGTTGCCGGAAGCGTTAGCGACGTGCGTTTCAGCAGTCCCCAAGGGCAAGGGAGTCCCGTACTCTGACTGAC
 TCGGCGAGGGCCGCGGAGATCACTCTGCGCCAGTTTGAAGACGGAGTACCAGGCTACAGCTCAGTCCGCCACCATTTGACAGT
 CTGGTCTTAAGCGGCGGTGGTGCCAAAGGTGCGGCATACCCGGAGCAATGCTGGCGCTAGAGAGAAAGGCATGCTCGATGGC
 ATCCGCGCTTGGCGGTTCCGTTCCGCTGCGGCTACCCGCGGCTTTTGGCCTCAGGTATGAGAGCAGCTTGGCTGCAAGTGGC
 CTTTCCGACAAGATGGATCTTATTTGCTGCTCGACAGCTCGAACAGAAGCTGAAGCTGTTCCAACACATTAGCAGCGAGATC
 GCGCATCGCTGAAAAAGGGCTTGGGCAACAAGATCGCGGCTTCTCTGAGTTGCTGCTCAATGTACTCCACGCGATAGATTGCG
 CGGGCTGAGCCCTAGAACGCTATTGCGCGACGAGACGCAAGGCGCTGCTCGGACAGATCGCTACGATCCAGAGTTGCA
 CGCCAGCGACGCTGCGGCGCATCGCCAGCAGATTGCGTCCGCTCCGAGTCACTTTGGCGATCTAGATCGGTTGAGTGCT
 TACATTTCCCGAGATTAAGACGCTGAACATCAAGGTACGGCCATGTTTCGAGGGGCGTCCGCAATTAGTGGTGTCAATGCCAGC
 CACACACCGGATCTGGAGGTGCGCCAGGCGGCACATATCTCCGTTTCTTCCAGGAGTGTTCAGAGGTGAGTTGAGTGAT
 CAGCCGTACAGGCGCGGCTAGAGTGGACAGAATTCAGGATGGCGGGTGATGATTACGTCGCGTCCCTGAGATGATCGAC
 AAGAAATTTGACAGCGGGCCACTGCGGCGCAACGCAACCTGATCCTTGAAGTTTCGAGGGCGAAGCTGGGAGGTAGCGCCGAC
 CGAGGTAAGTGGGGCGGCGCTCAAGGGCTGGTCTGCGGGTGCTGCGCTGCGAGCGCGCAAAATGCTGCAGCTCGAGGGC
 CTGGAGGAATGCGCGAGCAAAACGTTGTGGTGGCTTGAAGAGCGAGCGGCTGATTCAGTGGCATGCTCGGTGGCACCTTG
 AACTTACCATGCGCGACGAGATCAAGGCGCATCTTCAGGAGCGCTCCAGGAGCGAGTCCGTGAACATCTGGAGAAAAGCTCTT
 CAGGCTTCAGAGCTCATACCTTCGCTTCTCTCGACGAGGCGCTCTGGCACTTGATGACAGTATGCTCACCAGTCTTGTCTCAA
 CAGAAACCCGAGATCACAGACGGGGCGGTGGCTTTTCCGCAAGGCGCGGATGCGTTTACCGAGCTGACTGTCTGCTATCGTT
 AGCGCAATGGCTTGGCGGGTAGGCTCAAGTTGGACGAGGCTATGCGCTCCGCTCTTACGCGACTCGATGCGCTGGCAGATACT
 CCGGAACGCCTAGCATGGTTGGCAGCTGAGTTGAACATGCTGATAACGTTGATCATCAGCAGTTACTCGATGCCATGCGCGGG
 CAGACGGTGACGTGCGCGGCTGCTGCGCGCTGCTTAGCAGAGGCGCAGCGCGCAAAAGTGGCGGTTATGCGCGAAACATTCTGT
 AAGGAAGTTATCTTCCCTCTCTGTATCGCCCTGGCCAGCGGATTCCAACGTAGCTCTGTTACGTCGGGCGGAGGAGCAGCTA
 CGGCATGCCACAGTCCGGCGGAAATCAATCAAGCGCTGAACGATATCGTCGACAACTACTCGGACAGAGGCTTCTGCGTTTC
 GGCAAAACCTTGAAGTTGAGTACCTGTTGAGATGGCTAAGGCTTGGCGGAATAAGGAGTTCAATGATTGATACATGGCTGGCAC
 AGTGGGGCTTGAAGTTTCCCTCGAGCAACGATGCAAGCGCTGAGCTGCGCTGCAACCGGCGAGAGGACCGGATAGGTTAGGAGCGCC
 TCGAGGGCGGTTGGCTTTTCTGCTGAGTGGGACTTGTGCTTTCAGGGTTACCGCTGGGTGTGATCTTGCAATTGTTACAAG
 TGAATCTCTCAATCTCATCTTGGCA CCGGTGAACTTGCAGCGGACGATGCCGGTAGACTTGTGCTCTGGGCTGAGGCACGTG
 ATGGCGTTGACGATGTGGATGCACTGAACCGCTTGACAGTAGGCTGCGGGAAGGACATTACGATTAGTGCCATTGCTAGAGC
 CCACGGGTGAGTTGGTTCCAGCTCAGATACAAACAGCGCTGAGTGTGCTTGAAGCGCAAGGGGAAGGAGGATAGTGCACAA
 TAAATCCGCAAGTTGGTAATTTCTCTAATAGATGGGCCACCGAAGGTGGCTATCTGTTTCATGCTGCAATGAGTTGTGCTGCC
 AACACCATTTCTACCGAGCCCCATCTGGTTCAACGAGCCGAGGCCCTCGAGGTGGGCTTTTCTGTTCTGGAGTCCGGATATC
 TGATCCACAGCGCGCCAGAACGAGCGGCTGGTTTCTGCTTTCGCCAGTTATATCGGCACTGGCCACATTCTGTGCTGCCA
 GTCTGTTTCCAT

3/118
FIG. 2A

The Big Island: 84830bp (SEQ ID NO: 2)

CACCGGCGATGCATGGGATGCCGTAGGGGTCGCTGATCCAGCCAGGCGCCCTGGATGGTTTTCTGGTCGGCGCTGGCGGTCT
GGTTGTTGTTGCTCTGGTTGTCAATCACCTCCTCGGCCGCGCCGGGAAGGTGCGCACGGTCCCGTCGTTGAACACGAACGTGA
GGCTGCGGATCTGCCACGCACGAGGAGAGTGTCCAGTCGCCCCAGGCGGTCCCGCTGGCTACCGCGCCGCGACGTCCGGCA
GCTCGATGCCGTTGGCGGTGAGGTTGTCCGGGCCGATGAGGATTTTGAACGGGTAAAGGATCATTGAACCGTCCCGTCGACCGGCA
CACGACCGATCAGCGCAAAACATGGCCACCGAGCCCATGAGCGTCGAGTTCTGCGGCAGCGTGTAGGTCTTACGGACGAGCTTGC
GGTTTTCCAGGTGAGTCGCTGTTGGCCGACGGGGTGACGCCGTATCGATCCGCTCCAGCGCGTCTGTCCGCGATCGACCG
CATTGCCGAAGGAGTTCGGGAAGCTGAATCCGCTCGGTTGGGTGGTGAGCCGCGCGCCAGCGGCTGGCCATTGGCATCAACCG
CGCGGCGTCTGGGGCTCGATCCAGACGATATCCGATGAAGACGAGCCCGCTCCCTGAAAGTGCTGGCCATCCTTTGGTCGCA
CGCCGAATCCCTACGGGCAGGTGCGGTGTACCGGCCCTCGGAATGCGGTAGGTTCTCGAGCCGTTTCTGTACCTGGTCCAGGA
TCGTCTGGCTACGGTTTTCTGTCTGTGTGCTGCGATTCTGGGCGGTGTTGTTGAGCTTCTGCTCGATGTTCTGATCGATGTTGC
GCAGGCGGCCCTGCAAGTTCTCATTTGGCGGCTTTGAGCGAGTCATTCTCCTGGATCACCTTGTGATCTGGTCTCTGAGCTGCC
GGCTTTCCGCCACGATGGTGCGTAGTGTGTGCGCGGCTGTCCGCTCGATGCCAGGGTTGCCGCTTCTTCCGACGTTACCG
TCGGGTTGTCGGCGGCTTGTGACTGCGCATTTCTTTCTGCCCCATGATGCTCACGCCGATGCGGATGGCGGCGGCGGCTGGA
GGACGACCAGCAATTTAGGAGGGGATTACCGGTGAGCGCCACGACGGCCTCCTTTGGGATCGATCTGGCTGATGGAGGAGGGC
AGGAGCGCGTCGGCAAGGCGCGCGCGCGCTCACCAGATACACGCTAGTGGTGTGCGAGGCGTCGCCCCGGGCGCCCAAGTAC
GGGTGCTGGAAGGTGCGGCGCAGCAAAATGGCCATCAGGTCCCTGGGATCCAGGGCCAGGTGCTGGGCGCTGGCGTTCTGCAGC
TTCTGGTGGGGCGGCGCGCTCCTCAAGCTCGATACGTTGAGGCGTGTGCGAGTAGCCGATGCGCTGAGGCGCGAGCGGCTGGA
AGGTGAGCTGTGCTTACGCGCACCTGACCGACGCCATCCACCGGTTCCACCGTGCAGAGCGGGGATAGAGCATCTGCGCC
GCATAGCGCGTCAGAACACGGGGACGGGCGTTCGCGCGGCACGGCCTCGGTGCTTCTGCGTGTCTCGGTCTGTTTCTGCTGTCT
GCCGATGGCTGGGCTTCCCGGACTGGCCATAATGCGGATCCCATGGCTCGCCGCGCAGCATCCTGACCGGCTCGCGCGGTGT
TGGTGGCGGCTGCTTCTGCTGGCGCGCATATCGATGAGCATCTGCTGCCATTGGTTCGCGTCTGAGGCGCGAGCGGCTGGA
GGAATCGGCTCGTTGGCGAGCAGGTAGAGTGCGCCGCGGTACTCTGACGCGCAGCTTGCCCTGCAGATCCCGAGGAACCCCA
ACTCGCACGTTTCTGTCGACGAAACAAATGCGTTCTTGGCCGACCGTCAATGGAATGGCCAAACGGAATGCGTCCACGCGCAGA
ATCTCCACCGCTGGGCCAGTGTGGGTAGGGCAAGCATTAGCAAGAGCGAGCCTGTGCACTTCCGGATCATAGACCTCCCTCCC
GCTTGGTGGGGCGGCGCGCTCCTCAAGCTCGATACGTTGAGGCGTGTGCGAGTAGCAGTCCCACTGAGGCGGAGGAGGATTGG
TTTCCGGGTCGACGTGCGCGCGGATAACGTGCAACGGGTAGCGGGCCAGCGCCGCTTGATCTTCTCGCCGCGTAATACTCCG
TGCTGTCCATGTCCAAGTTGACGCTCCAGTCATTGATCGAGTGTGATCACGCGGCCATTGCTCTCGCCAAATGCCTCGACCGG
GGATTTCCGAGGTGGTGCCTCGCGACCCCTGAGCTCGCGCGGTACGACGAACTCAAAGTCTTTCTCCAGGAAGACTTTGC
AGGAGGAGTGAGGTAGGCAGCGTAGCGGAACAGGTTTCTTGTAGTCCACCTCGCCGCTCCTTGGGCCAACGCTGCACCTGCT
GGAAGATGTAGAGGCCGAACGCATAGACGCTCTCTGGCGGAATGTCCCAACCAACCGGTGCTTCTGAGCGCAAGTCGGGCG
GAATGTGATCCAGAGGTCTTACGGGCTTCCAGACGCAAGTAGGCCAGCACAAAGATGACCATGCAGGAAGCCGGTGATGA
ACCGGAAGGTGTTGATGTGTGCTGCTGTTGCGCAGTGTGTTTTCTGAAACTCATGGGGACTCCGTGCGTGCAGTGGTCCAGGC
TCCGTGATAAGTCGAAGGCTACGCGGTTGGCGCGCGGTGAGGCGCATGTAGTCTTGAGGCTTGGCGGATGAGTGTCTCGATGGGTGAGAA
ATCGGGCGCCCCGCTTATCCGCCGACGAGCGCGCTGCCGATACCTAGGCCGAGGGCGCCGACAGCAGCGCGCCAGTGG
AATGCAGGCAGCGTTACCTGCGCAACAGGCGCGGGATGCCAAGAACGAACCCGCTGCTCCGCTGGTGAAGACCGTGATCCA
CATTTCGTCTGCGGCTCAGGCGCGCGATGACTACCGGTTCCCGGTTCAAACGGGTGCGCAGGAAGCTGAGGGTTCCATCTTGAA
CAGATGTTCTTCTGCGGATGGGAAGGCCCTTACATGATGGCGGTGCTTGGTGACGAGATAAATGATGATGAGTACGAGCGCA
CACCTACGGCTACGCCCCGCTCCGAGATCCGACCACTTCTTCTTCCGTCATGGATGGCGTGATAGGTGCCGTAGGTATGCCAAG
CGACCCCGAGAAAGACAGCCGCGCAGATGAGCAGCGCAGGAGCATCGCTCCGTCATAGCCGAAGTTCTGGATGGTTTGATGA
TGCCCGATCCCTCCCCACGGCTAGGTGCCTCGGGTTTGGGGAGTGACAGCAAGGCGATACCGGGGAGCGCCAAAGTGTGGCGG
CCAGGCTCTGGCAGGGCGGACAGTTTCTGGAGGTGATGCTTCCGAGTGTGCTCCGATGAGTGTCTCCGATGGGTGAGGAGAA
AGAGGTGATGCGGAGGAGGACCAAGATCCGGGATCGCGGAAGCGCCACCGGTGGCGCTGTGCAAGGTTGTTGGTGGCCAGCCG
CGCCAGGTGCTGTACATCGCCAGGCGGACACAGCAGAAGGAAGGTATCGCTGCTCCAATGAACAGTCCCTCGCCGCGCGAT
GGGGGAAAGCCAGCGGCGGCTGGAAACGCTGATGTCTGGGCTCCGCTCATGCTCATGGCATCGGCCCTCCGCGCAGGGTGTAAT
TCCCTGATAAGTCGAAGGCTACGCGGTTGGCGCGCGGTGAGGCGCATGTAGTCTTGAGGCTTGGCGGATGCGCTGCAGGT
CAGCGGCGCAGGCGCGGGTAGTCGAAGTAGAAGCGCTGTCCGGGCTCATCGGCGCCCTGGGCACTGCGGCGGGCGGTGTCTCGA
GGGCGTTGAGCTGCCGATCATCACTCCAGGTTGCGCTGCTCGGAGGCGCTCGTGCAGAAAGTCCCTGTGCGACAGCGCAC
TGCAGGCCAAGACGCGCCAGGACGCTTCCGCGAAGGGATGATGTGGAGTGTTACGGCGAATCTCGTGCAGATCAGGGTGCGAT
CAGCTTCCGATCGGCTGATTCCATTCCGCAAGAAATGAAGGCGCGGTCCGGGTATTTTTTATGAGGTTGCGCGCTTGA
AAAACAGCAGGTAGCTGACGGGACAGCGACGCCCCAAGCCAGCCTTGGGCGTTTTGATTACCGATCTCGGTTACGCCCGCA
TGAGTGGCTGAACGCTACACAGTCCAGCCTTCTTGGCCAGACGACGACCTCGGCCAGCGCAGAGGGTTGAGCAAGGCGT
GGTCGTGCGGTTCCAAGGTGTCTTCTTGTGTTTGAACATCGCGACGCCCTTGGTGAGGGTTTTGTAGGGAATGACGATGTTCT
CGACATGGTATTTCGAAGGGCTGTTTATGCTCATGGTGTGCTTCCGATCGGTTGAGTAGAAGCCGACCGTTTCAAGACAGG
AGGAGGGCCTCAACCGAAAATACTAGGCCTAGCACCAAGTGAATTTCTGGTGGTTGGCGGGTTTAGAGACACGTATGGACGCTG
GGTGTATATTTTATATATATTTCTAAGGGAGACGGCTGATGCTTAGAACAATCTTATTGGAGTTTGTAGCCATGGCTGCTA
TGTTGGGCGAGTTATGGGGTGGCTGCCGCTACATTACGATGCGGGTCGGCAATTGTTAGTGAGGCGCACTTGATGATGATGTC
TTAGAAAGTCGGCAACCTGATAGCGGTAATAATTGAAGGCGCGCATGGATGGTAGTGGCTATAGTGGCGGGGCTGCTA
CTGTGCAAAACTGGGTATATGGACCAAGGAATGGATGGTACCAGAAGCTTAGGTTTGTGATGGAAAGACTAGTTCAGATAAAAG
GCAGTATGGACTAGGGTATAGCCGTGGATGGTGTGTTTTCATCCACGGCTATAAGTCTCATCCGCAGATGATATAAGGGAAG
GATATTTGCGATTTGGTAGGCCTTGTGCGTCCGAAATAAACACGGTTTGTACTGGCGCCTGCACGGGAAAATATCTGTTGAGGT
TGTTCCGATTAGACTGCCACCGTTGTAATTGGCTTGGAAATGCTTGTGGAACCTGCTATGTAAGTTCAAGTAGAGTGC
TTTGGTTGGCGGATTGTGCGGAGTTTATGCTGATTGATGTGTTTGTCTGCTCGCTCGGAGCGCCTGCTGGATGTTGTTGAAG
TGAGTTGGGCACCATTAAGGTTGGCGCTGAATTTTTTAGTGTGCGCGCATGTTGATGGCCTGTAAAGGGTGGATATATTCG

CCGCTGAGCGGTGTTGACGCGATATTGGTACCCGAGGCCGACGCGTGTATCGCTTGCTGAAGGTTGTCTGCTGTTAGTTGTG
CACCATTGAAGTCGGCGCTAAACCTTTTTCGTGTCTTTTGCAGAGTTGATGAGGCTAGTAGGGTGCTTATTATTAACCGCCTGTG
CGGTATTGACGCTAATGGATGTTCCGGTCTCTGCGCGCTCTAAAGCTTGTGAATGTTTGAAGTAAGTTGAGCGCCATTGA
ACTCTGCGCTGAATGTTTTTGTGTCACTCTGTCATGAATAGTTTGAAGAAGCGGGTTATATTGGCCGCTGTGCGGATTGA
CGCTTATGCTGGTGTTTTGTCTGCGGCATTTACTGCTCTAAGTAGGTTGTGCTGCTGAGTTGGGCTCCATTGAATTTCTGCG
TAAAATTTCTTTGAGTTGCCTGCTGTAGATATTGCGGAAGAAGTGTGTTTATATTAATGGCTTGTGCGGCTATTGTGGATACT
TGGTTAGTGATCTCTGCTGAGTTTATGACTGTGAAGTGTGTTAGCAGCTCAGCTCTGCGCGCTTGAAGATTGTTGAGAGGTTTT
TTGTGTTTTCTGTCAGTGGTTGTTGTCCTTGAGAAGATTTCTATATTAAATGCTTGCCTGCTGCTTATGCTGATGCTGTGATTTT
TTCTGCTTTTGATAGGGCTCTAGCATGTTGCTTGGATTTAGAAGGCGTCCGTCGAAATCAAGGATAAGTTTCATCTCAATGC
TCCGCTGTGACCTAACGTTAAAGTCTTGGTGTGAGTACTTCTAGGCTGAATGAGTTGGCTACAAATCTCTTTGCTCTTATGTCA
ATATCAGGCATTGCTGTGACCAATATCATTTTTTTGTAGTCAATTTTTACAGTAGCTATGCTCTTCACTGTTTCTGTGCTCC
ACGTACCCTGCATGACGACGACCACTAGGTTGTTGATGCGAAGGATTTTCATGCGCGGGTCACTGCGACGTAGATCAGGTTT
AACTCATCGTCGCGCTTGCCTGGGTGGTGTGCGGGGCCAGCGGGTCCGCGTTGAAGTCGTCGTACAGGCAACGAAATCCCAT
TCCAGCCCCCTTGGCCTTGTGTGCGGTGGTCAGGGTGATTTGTTGCGTCCAGCTCATCGTCAAGGGTCAGTGAGCGAAGCTCAAGG
ATCCGCGCAGGCAGCATCAGGGTAGGTCGATATGATCTTGTAGCAGCGAAGCATCTCACCGTCTGGTGTATCTCGGCGATCTCC
ACGCTACTGGGTGTAGTCGCGGTAGTCACGAGCAGTTTCTGTTCTGAGAGTTTGGCGCAGGCTCGGCTGAATGCGTACAGA
TCCTCCAGGTCGCGCAGCAGTAAGTGTGATACCGCTACCCAGTGGAATTTGGGCTCCGGATGATTGCGGACAGCTGCGAGG
GCATTCTCGATGACGCGGATAACGGTGTGCGGTGAATGAAGTGTGCGGTGAGGAAGGTCCGCGCGGAGGGAATTTTTACACGCGT
TCCGGAACCCAGCTCTTGAGTAGTTTCCGTGTTTCCGCTTTGTAGGAGGATGATGTTGGCCAGTGTGCGATCGCGGCGGCAAT
CGCGAGCTCTGGGTCAAGTGTCTCTCGCGCGCGGCTCCAGTGCAGTGTTCAGGGCATCTTCTGCGCCCTGAACCGGTAG
AGCTGCTGATGGGGATCGCCGACGATAGCCATTCTGATGCGCTGCCAATGGGCAATGTCCGCGATCACTGGGTTGATGTCTTGC
CCCTCGTCCAGGAGATGCACTGCAAGCGCTGGCTCAATCGGGCTTGTCTCAGTTGATACAGCTTCAGGTAGCCGTCAAGGGGC
ATCAGCATGCGCGTGTCTGGAGATGACCATGCGCTCCAGTCCAGTCTCGCGCATGTCACGGCCCTGCTTGAGGAAGCGCTTC
TGAGCATGGGTGAGGAACGCTTGTGTCGGAAGCGCGGAAGTGCCTGCGCGAGTTCCGCGTCGCGGCTGGGCACTGATGTT
TTCAGCGTGGCCAGCAGTCACGTACCAACTCCAGTCTTGGGTATCGAGTCCGCGGCGGATATCGGTGAGTGTGCGGTTCTTCT
GTCTTCTGTGGGCGTACTGAAATGCCGTACACGATAGACCCAGACTGTGGGCGGCTTGCACACTACGTTGCGGGGAAACTTG
CCCTTCGCGGCTTTCTCCACGAGCTGTTGTAGCAGAGATAGAGGATTCTCAGGCTAGGTTCTGACCTGGCAAGCCACAGG
GTGGTAGTTTGTGCCAGTGCTGCAAGGCTGCGACCAGTACTTCCGTGCTTGCATGATGATCGTGTGACTGTTGCTGAGTC
CACTGCACGCGGGTCTCCTACAGGTACTTTTTGAAGGTGGATGCTGTGATGCATACCGCTACGCCGAGCAGCGCGGCGCAGGG
CAACAGGATGAGCAGGGGATGATGCTGATGGGGATTGCCAGGTAGAGCGTCCAAGGGACGACCGCTAGCGGAATGATGCTGCC
GCGCGCCTTGTGGTAGAGGTAGCTGGAATCCCGCGCGCGCGGCAACTTGCAGAGTCCCGCGCAGCAGGCGCTCCACAGGCG
GGTGAACCGCGCCATCAAGAAACCGGATGCTGATGACAGGATCAACGAGGCGACAGCAATGTGACGACCGGTGACAGCG
GGCCAGGCGGTAGTTCTGACGCTGGAGACACCGTGGGCGGTGAGATGCGAACGTCACGCGGGCTCGTGGCCGCGGCTTGC
GATGGTAGTTCATGTTGGTTCATCCAGTCGACCATCCCGGTCTTACGAAACACCAAGTCATAGGCAACTGGGCCAGCCAGGTGGC
GGTTCGACCTGGTCTCTGACGACGAGCTGTGTAGCAGCCCTCGCGACGCAACTGAGTTTCGTACTCAAACTATGGCTTGGG
ATGCTTCCAGCCGCGGTGAGGCGAGAAATACAGGATACCCACTCCAGTATGTCGCCGATCAGCGACCCGAACATCAC
GCCGATTAGGAGCAGGAGTAGGGGGGCAATCGAAAAATGCCAAGACTGGTGGCGTGGGTTTTCTGGGTAGCCGCGGAGAAATG
AACAAGATGTTCAACAGTTCTCTATAGGAAATCTGTGTTGTAACCTCATAGAACTCCCTTTTACAGATGACAGTTGTTGCTT
GGAGTGTGTATCCACTAGTTCTCGGGGAGGCGGTCACTGTAGATTTTTTGGGGGGGTATCTCGGCAATAGGGGGGATGAATAGC
GTGTGCTGCTGAGCTCTGAGACCTATAGGAATCATCTGATAGAGGGAATTTTCCATGCGCTCAGTCTTTCCGATGAGATT
GAAGAATTTCTTACAGCTTTTTGATAGCGTTTTCTCACTCCGAGTGCTGCGCTCAAGTGTCTCCGATGCTCGGTGGCGCTT
GATGTTGTGCGTGTATGACGCTGATATTAGTGTGAGCATGGTTTTAAACCATCAGGTGTCTTCTCAGTCGGCAAGCTATGAA
TGTGGCTATGATCAAGAGCGCAGCTTTATTTGAGCAGCGCGAGCGGCTTGTCTCAATCAGTTGAGAGGCGCAATGCTGTCCTTGGC
CGCGGTGAGAGCGCTCGTCAACGAAGCGCCGAACAATGTGAGCATCTTGGCTGTAGTGACGAGGCGAGGTCTGCTATTGAC
CGCTCGACGCTCGGTGATCTCCGGGAAAGCGGCTGGCACTGATGATCTGTGTCATACGACAAAGGCCCTCTGGTTTACCG
GCTTACCGCGGATGTTAGGCCCTCGGCAGCGATATCCAGCACGATAACCAAGAGGTGTACCGAGCCTTGTGCGCACTCCGTC
GGCGCCTGTTCACTGGGTGATGACGGTGGTACCCTCAACGGCTGTACCTTTTGAATCTCTAGGCGATGAGCGGCGGAGGG
GTGGTAGGCTGGGATTTCTCGGCGAAGACTCGATTCGATGTTGCGCGGAATGATGCCGGAATACATGCTGTGCTGATCA
GCATGGGCAAGTGTGATCTCGTACGGAACGACGAGGCGCTGGGAGCGGTGCGTGCAGGACGCTTTTGTGAGAGCGGCTTCCG
TTTCATCGGTGCTGGCCACTGCCGACGATATGGTGCTTTTCCAGCACGTGGGGTCTTTCAGCTGGGATCTGATCTATCACA
CGGTATCGGTGCGCTTGTGCTGCTCTGTGCTCTCTTCTGTTACTTGGCTTCTGCTTGGCACTCGCATCGGATCCTACTGCA
TTGGCTGGTGGCAGCATCGAGCACTGTTGATAGGCGCGCAAGCGCTTGAAGCATTAAGGAGAGCGAAGCCTTTTCT
CCGTGCACTTATCAGGCGCGCGCGCTGCGCTGTGCGTGTGCGTGCAGCGCGCAGTGGTCTTGGAAATCCCGAGGC
CGCCAAATGGCTGGGTGATAGCAGGCGATTGCCACGACGCGCGGAGATGGATTCCAGGCGGTCGAGGAGGTGTGAAGTG
TTCTGAGCAAGAACTGGAACCGAGGAGGCTACATCTTCATCTCAATTACAGCCCAACCCGCTATAACGGTGAAGACGTATT
GTTCTGCGCCTTCAGTGAATCAGTGCAGCAAGCGGATGGAGGCGGAATGGCTCGCGCAAAATCCCTGGCGGATGCTGCCAA
TGAAGCCAAAGACGTGTTTCTCGCCACCATGAGGCATGAATCCGACACCTCTGTACGCGATGCTTGGACAGCTTGTGAGTGTCT
TGGGCGTACCGAGCTGAGTCGGCAGCAGGCGGTTACCTAAAGGCAATCCAGCATTTCTCGTCGACCTGCTGCAACTGATCAG
CGATGTGCTTGAAGTATCCAAGATAGAGGCGCGCAACTGGAACCTAGAGTGTGCTGGAATTTCTCCCGCTGGAATTTGACCGAAG
GGTCTGTCAGTCTGTTACCGGTGACGCGCGAGGCGGCTGCACTGTATACCTGCTCTCTGCGGAGTCCGCTGCGCAT
CGCGGGGCGCGCGGCTGATCCGCGAGATTCTCAACAACTGCTGAGCAACGCGTGAAGTTACCGCAATGGCTATGTCAA
CGTCCACCTGAAGGCCAGCGTGGTGCATGCCGAATGTGTGATGCTGACCTGGCAGGTCAACGATACCGGCATGGGGATCAACGT
CGAGGATCAGCGCGCTCTGTTGCAACCGTTCTACAGATACGCGCTCCGAGCATCCGCTGCGAGCATCGGCGCTCGGCTGTGCT
GATCAGCCAGCGCTTGGCCAGCTAATGAATGCGACTCTGAACCTGGTCAGTGTGAGTGGGTCAGGCGAGCTTTAGCTCAG
GCTTCCGCTTGTAGCGGATCGCATGAGGTCAGGCGCAGGACCTAGCCGGGTGCGCGCTTCAAGTGTCTGGCGCTGTCTCGCA
CGTCCACCTGAAGGCCAGCGTGGTGCATGCCGAATGTGTGATGCTGACCTGGCAGGTCAACGATACCGGCATGGGGATCAACGT
CGAGGATCAGCGCGCTCTGTTGCAACCGTTCTACAGATACGCGCTCCGAGCATCCGCTGCGAGCATCGGCGCTCGGCTGTGCT
GATCAGCCAGCGCTTGGCCAGCTAATGAATGCGACTCTGAACCTGGTCAGTGTGAGTGGGTCAGGCGAGCTTTAGCTCAG
GCTTCCGCTTGTAGCGGATCGCATGAGGTCAGGCGCAGGACCTAGCCGGGTGCGCGCTTCAAGTGTCTGGCGCTGTCTCGCA
CGTCCACCTGAAGGCCAGCGTGGTGCATGCCGAATGTGTGATGCTGACCTGGCAGGTCAACGATACCGGCATGGGGATCAACGT
CGAGGATCAGCGCGCTCTGTTGCAACCGTTCTACAGATACGCGCTCCGAGCATCCGCTGCGAGCATCGGCGCTCGGCTGTGCT
GATCAGCCAGCGCTTGGCCAGCTAATGAATGCGACTCTGAACCTGGTCAGTGTGAGTGGGTCAGGCGAGCTTTAGCTCAG
GCTTCCGCTTGTAGCGGATCGCATGAGGTCAGGCGCAGGACCTAGCCGGGTGCGCGCTTCAAGTGTCTGGCGCTGTCTCGCA
CGTCCACCTGAAGGCCAGCGTGGTGCATGCCGAATGTGTGATGCTGACCTGGCAGGTCAACGATACCGGCATGGGGATCAACGT
CGAGGATCAGCGCGCTCTGTTGCAACCGTTCTACAGATACGCGCTCCGAGCATCCGCTGCGAGCATCGGCGCTCGGCTGTGCT
GATCAGCCAGCGCTTGGCCAGCTAATGAATGCGACTCTGAACCTGGTCAGTGTGAGTGGGTCAGGCGAGCTTTAGCTCAG
GCTTCCGCTTGTAGCGGATCGCATGAGGTCAGGCGCAGGACCTAGCCGGGTGCGCGCTTCAAGTGTCTGGCGCTGTCTCGCA
CGTCCACCTGAAGGCCAGCGTGGTGCATGCCGAATGTGTGATGCTGACCTGGCAGGTCAACGATACCGGCATGGGGATCAACGT
CGAGGATCAGCGCGCTCTGTTGCAACCGTTCTACAGATACGCGCTCCGAGCATCCGCTGCGAGCATCGGCGCTCGGCTGTGCT
GATCAGCCAGCGCTTGGCCAGCTAATGAATGCGACTCTGAACCTGGTCAGTGTGAGTGGGTCAGGCGAGCTTTAGCTCAG
GCTTCCGCTTGTAGCGGATCGCATGAGGTCAGGCGCAGGACCTAGCCGGGTGCGCGCTTCAAGTGTCTGGCGCTGTCTCGCA
CGTCCACCTGAAGGCCAGCGTGGTGCATGCCGAATGTGTGATGCTGACCTGGCAGGTCAACGATACCGGCATGGGGATCAACGT
CGAGGATCAGCGCGCTCTGTTGCAACCGTTCTACAGATACGCGCTCCGAGCATCCGCTGCGAGCATCGGCGCTCGGCTGTGCT
GATCAGCCAGCGCTTGGCCAGCTAATGAATGCGACTCTGAACCTGGTCAGTGTGAGTGGGTCAGGCGAGCTTTAGCTCAG
GCTTCCGCTTGTAGCGGATCGCATGAGGTCAGGCGCAGGACCTAGCCGGGTGCGCGCTTCAAGTGTCTGGCGCTGTCTCGCA
CGTCCACCTGAAGGCCAGCGTGGTGCATGCCGAATGTGTGATGCTGACCTGGCAGGTCAACGATACCGGCATGGGGATCAACGT
CGAGGATCAGCGCGCTCTGTTGCAACCGTTCTACAGATACGCGCTCCGAGCATCCG

5/118
FIG. 2C

CCTAACGGAATGCCTGTGTGGCTGGATCTCCGCTGGGGTGAAGGGCCATGGTCGCGACGCCGAGGTGCTGGACGAGGCCGA
 CGCGACCTCGCTGCTGGTCAAAGTGTACTGTGGAGGGGGCGCCGATGTTTGAAGCATGGCCAGGATGCCGGGTGGAGCTTTC
 CCCTCAGGGTGATATGGAGCCGACGGCAAGGGCCGCGACTGGCTGCTCGGGCTCAACAACCTGAACGGCTGCATCGTCTCT
 GGGCCTGGCCATGGCGCTCTCGCTGATCCTTCGACGCCGCCGATACGGCTGGCTCCGTTGCGCAATCTAGGTCTCCGCGTCT
 AGTGGTGGAGGATAACGCGATCAACAGTTGATCTTGAGGACACAGATGGAAGCGCTGGGCTGACGCTGGAGCTGCTCTTCGA
 TGGTCGCGAGGCGTTGCTGCACTGCCAGACGGCTGCTTCGACGCTGGTGTCTACCGATATCAACATGCCGAACATGAACGGATA
 CGAGCTAACCGCGGAGCTACGGCGCCAAGGGTTCGGCGACCGGATCATCGGCGCGACGGTGAACGCCATGCGTGAGGAGCGCGA
 GCGCTGCATGTCCGCGGGATGAACGATTGCTTGGTCAAACCGGTGGATCTGAATGCCCTTCAGAACTGCTTGATTAATATTCT
 CAAGTGGATCGATGAGCTGGAATCCTATCGGGTGTGGTGTGCAAGATCAGCCGTTTCAGCGCAATACCTGCTCAACCTG
 TTTTCGAGTACGCGCGCTGCAGTACCTAGGTGCGCGGCGACGGCGAGGCGTTGCGCTGCCGTGAAGCAGGACAGGTTTCGAC
 CTGATCCTCAGCGATCTGATGATGCCGGCATGGATGGTATCCAAATGATCCTGCAACTGCCGTATCTCAAGCATCGTCCGAAG
 CTGGCGCTGATGAGCTCCTCGTCGACGGATGATGCTCAGTGCCAGCGGGTTCGCCAGAGTCTCGGCTTGTCCGTTAATCGAC
 CTGTCGCCAAGCGACTCTGCCAAGGCCATCGGCCAATCTCTGGAACACCTGGAAAGATGCCCTCAGGCAGAGCTGGAGCCG
 GAAACCGACGAGACTCCGATGGGCGCACGGCGTTGCTGCTGATGCTGATTAACCGAGCAACTGGTACCTGGTTCAGGCTTAAG
 AAATCCCTCCACACCGGGCGCATAGTCGGCGCGAGGCGTTGATACGCTGGAGCCACCCGACGATGGCCTGTTGCTGCCCAGC
 TGTTTCATGAGTGATGTCGACGCTACCGGTTCGACGAGGCGTTGCTTGGCGCGTGTCTGAAACAGACCCCTGAAACGCCCAGGAA
 TCGTGGCGCAGGGCGGGTTACGAGATTCCGGTTTCGGTGAATCTGCCCGCGACCTGCTCGATAACAGGAACCTCCGGATCGA
 CTCTATGAGTACGTCGCGCTCGCGGGCTTGTACCGCTCACTATGTTTCGAGTTGACCGAGACGCTGCTCAACAGCTTCGCTCA
 AGTAATACTATGACAGGTGCTGTGCTTGGCATGAAAGGGTTCGATTGGCCAGGACGACTTTGGCCAGGGTTACAGCTCG
 TTCTATAACCTGGTACGACGCCCTTTCAGGAGCTGAAGATCGACCGCTCCCTAGTCCAGGATGCGTAGAGGATAACGGCCTC
 AATGCAGCTGTCTCAGTTGATTGAGTTGGGTCAACCGCTGAATCTCGACGTTGGTGGCCGAAGGCGTGAGACCTGCGAGGAA
 CTGAATCTTTCTCGTCTTGGCTGCGACCGGGCGAGGTTTCTGATTTCTAAGGACAGTGTCTGCTGAGTTGCTGAGCTGAGCG
 CAGTTAAGGGAGGACGGCCCCAGCCTCCTTGTAAACGCAGTATCCCCATTATCGCGGAGTCGATCGCAGCCAAACACCGTCAG
 CGCAACAGTGTGCTGAAGGAGGCGCATCCCGTGAAGTCTGCTAGTGCCCTTGGAGCACGACAAACAACTTTTGTCTCAAATGGA
 CAACCTCTCGCAGAGCCTGAGCATCGGCTTGATCTGTGGTGGTGTGCTGACCGTATTGCTGTTTACGATCTGTTACTGGTTCG
 TGGGAGATTGTTTTCAGGAGGAGGAGGACAAAGTCTCTTCCACTTACCCGATGATGGATGTTATACCGGAGCATGAGGTAT
 TTCTTGGGCGCATCGCTCGAAAAGCGACAAGACCACCCAGAAGTACGACTATGACGTTGGTGCCTTTGACGCGGCACCTTGTGG
 CAAAGGAAAACCGATTAGCGGTCTATGAGGAGCGGGAGTTTCTTGTCTATGCCATTTCTACTGGCTACCAAGCACGCGTTGA
 GCGCCGATTCCTCGGAGATCCGTTTTCGCTCGGTGATTTGCTCGCCAAATTTCTACGGAAGCTTCTGGAGTGTTCGCGCTATC
 CCGCGCCACAGTTAGTATCTTTGATCTTTCCGCGAGTACCCCGCTGCGAGTGCCTGATTCCCTCCACAGCGCAGCGTGACA
 GGTGAGCGGAAGCTATCCGATGATAGTCGAGCGCATTCTGGCGCGCTTGGCACCCCGCGGTGGGGGAGGACGCTCAGCGTG
 TCCATTGGATACGCGCTGATCGCTATCGCGACTCGGCGCTGGAGATGTTGGGAGTGCCTCGGCTGATCTGCCGGAAACACTCT
 GGTGGCACGACGAGCCGAACATCTGATCATCGCTGCGAGCCTGCTTGTATCTCAGGCGAATCAATGACTTCGAAACAGTTGGTGG
 AGCGCCCGCATTCGATTCGTACAGCCTGGTATCGCGCGCTGGTATGCTCGGCGCGCCCTGAGTATTGCTCGGCGCGCCCTGAGG
 ATGGCCTGAACCTCACCCGACAGGGGGTTCGCGTTCAACTGCTCAGCCAGCCTGAGAACGTTGGCTCGCGGTCTACCGAACCG
 ACTACGGCAATTTCTTTCGCCACTCCCGGTGGCTGGTGGCAGGTCTGCTGCTGACCCCGCGCTGCTCCTGGCCGCTTGGCTCG
 GGATGCGTTGGTACACGACGAGCCTGCTCAACCCGCTGATCGGCGGCAACCGCAACTGGTGGAGAGCGACACCTTCAGCCGGA
 CGTGATACAGTACGCGCGCTGCTTGGTGGTGTGCTGACCCGAGTACGAGCAACTGGTGAACCTGGAGTTGCTCGACTGCGGCA
 AGTGGCTGGCGGGCCACGAGATCCTTGGGCTGACTTCCAACCTGGAAGCTTTTCGATGCGCGTGGGCGAGTACCAGGAGACA
 TCTGTATCCAGGTGCGTGGGCGCTATTTGACAGACCGCTTCGCGCGGACCCGCTATGCCGGCACCGAGGCGGTACTGTGCGTAT
 TCAACGATCATCAGGTCCACTGCGAGGCGGAGACCGCGCTGCTCAATGCGAAGCGAGCAGCGGATGCCCGCAGCCAGGCCAAG
 CCCTGTTCTGGCCGCGCATGAGCCATGAATCCGATCCTCCCTGCTGAGTGTCTTGGCACCTGGAGTTGCTCGACTGCGGCA
 CCCTGAACGAGCGGCAACGCGCTACCTACGCAACATCCAGAGTTCTGCTGCGACGCTCATGCAACTGATTAGCGATGTGCTGG
 ATGTCTCGAAGATCGAAGCGGGGCGATGGCTCTGACCTGGCCGCGCTTCAATCCGCTGGACCTAGTGGCGGAAGTGTGTTGGCA
 ACTTTGCCCGCAGCGCATGGCCAAGGACCTGCAGTTCTATGCTGCTGACACCCGAAGTGGCGGCGCAACTGATCGGTGACG
 TGACGCGGATTCGCCAGGTGCTCAATAAATTGGTGAATAACGCTGAGGTTTACCGATATCGGACGCGGTGCTCCTGCGCGTGA
 AGTTGCTCTCCCGCAATGATGGTGCAGCCCTGTTGAGTGGCAGGTGCGCGACACCGGTATCGGTATCGCACACGAACAGCAGG
 AGCGCTTGTTCGAGGCGTTCTACAGGTTTCGGGAGCGCACCATGCCGCGCGCACGGGGCTAGGACTGTGCTGATCTGCTGGCATC
 TGGCGGAAATGATGGGCGGTCACTGCGAATGGTCAGCGAGACAGGGCTCGGCAGCAGCTTCAGCCTGGTGTGCTGAGTTGCCCG
 AGGACGAACAGTCCCGGCTGGCTTGCCGCGCGGGGCTTGAATAACCGCTTGGCTCCATGTGCGCTGCCCGTGGGAGCTAG
 CCGACAGCGTAGGGGCGTGGCTGAAAGCCTGGGGCTGCAAGGTGACGAGCGGCGAGGCGCGCCCTCCGAGCTGGAGACTTGTG
 TGCTTCTGGAGCTGTGCGGATGGCGGCGGGCCTGCTTCTTCGCCCTGGCCAGGCCCCGGGTGCGCGCTCCATGGATGCGC
 CTTGCCAGCCGAGCTGCGTGAGGACGGCTGGCGTGTGCGCTGCACAACCTGGCGGGAATCGGCCAGGCGCTGGCGCAGGCTC
 TGGGTGGCGATATCCCGGAGCAACGCGCGGCAATGCCCTGCGCGCGCTCGGGGAGACTCGACTGGAAGTGTGCTGCTGCCGAGG
 ACAACCCAGTCAACAGGCGCTGCTTCGCGAGCAACTGGAAGAGCTGGGTGTGCGCTGAGCCTTGGCGGCGATGGGCGGCGAGG
 CCCTGCACTGTTCCGACAGTGGTCTGCTTCGACCTCCTGCTCAGCGAGCTCAACATGCCGAACATGACCGGTACGAACCTGACCC
 AGGCGCTGCGCGAACGAGGCGAGACGCTGCCGATCATCGGCGTGACCCGCAACCGCCTGCGAGAAGAGGCGAGCGCTGCCGGG
 CAGTGGGAATGCAAGATTGGCTGGTGAAGCCGATCACTGCTGATGCTGATGATGCTGCTGAGTGTGCTGCTGCTGCTGCTGCTG
 TCGTGTCTCCCGCGAAGCGCGAGACCTCGGCGCGCGCGCAGCTCGACGACGGTCTCTACCGCAGGTGCCGGAACGCGATGC
 GCGCGCTTTCTTCTGAGACCATGGGCAAGGACCTGGAGGCGCGCGGCAAGCGATTTCGCGCGCAACGACCCGAAGGGGTGCGAGC
 AGGACCTGCTGCTGATGGCGGCTCCTTGGCGGTGATGCTGCGCGCAACGCTGGTGGTGTGTCAGGCGCGCGGAGGAAGGCC
 TGCTGGAGTTCGCGCTTGAATGTTCCGCGTGGAGATTGGCGAGGCTGCTCGTTTCATATCGAGCAGGCGCTGGAGTTTGTGAGAA
 AGACGGGTGATGCTGCTCGGGGTCAATAGTTGGAATTAGAACAAGAGAAGCCTATGAGTAAGCTCAAGATAGTACTGGCCGA
 TGACCATCCGATCGTGGTATGGGCGTATGCGACATGCTGAGCGCGACGGTTCGAGGTGGTGGGCGAGGCGCTCCACGCC
 CAGCGAACTGGTCGAGGTGTGCCGCGAGGCGAGCCGATATCGCCATTACCGACTACAGCATGCCCGGGGACGAGCGCTACGG

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND
 PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September 12, 2003 Serial No.: Not Yet Assigned

Page 5 of 118

Customer No.: 21559

FIG. 2D

CGATGGACTGAACTGATCGACTACCTGTTGCGCAACTTCTCTGCTACTAAGGTGCTCATCTTACCATGGTCGGCAACCGCCT
 GATCCTCGACAGCCTCTACGATCACGGGGTGTCCGGCGTGGTGTGAAGAGCGGCGAACTCGACGAGCTGCTCTTGGCGCTCGA
 CGTGGTGAAGCAGAACCGCGTCTACCGGGGCGCGAACATGCTCGACCCGACAGTGTCTGGCGAAACCGCGACGAAGTGAAAG
 CCGCTTCGCGCGCTTGTGATGAAGGAGTTCGAAGTACTCCGTCACTTCGTTTCCGGCAGCAACGTCTGCGATATCGCACGGCT
 GCTGAAACGTAGCGTCAAGACCGTAAGCACGAGAGGTCTCGGCGATGCGCAAGCTGGAAGTGAACAGCGACCCAGGCCTTGAT
 GACCTTCTGCGTGCATGCCAACTTGTTCCATTGACAAGTTTCCACCGCGAGCGTTCGAGACGCGCGCGGCGATGTCTCAGTGA
 AGCAGGCTCCGTCCCTGCTGGAGATTGCGCGCGTCACTGATCCAGCTGAAGCGCACCTCGGCCTGGCCGCGGGCAGCGCCTGG
 CGGAGCGCCAGGACGCGCCTGGAGAAGGGGGCAACAGTTTCGGCGCGCGCCAGGCGGGTGTGGCCGCCAAGCTCGGCGCGCGGCG
 AGGCTGTAGAAATAGGGAGTAGGGTTCTCCATCTGTAGTTTTCGCGCCGCGCTCGAACCTGAGCTGGTGTGGCGCTTCTTCC
 GCACCTTGGCCAGAGGCGAGGCGGCTAGATGACTTTCATCTGGGTGCGCAGTGTAAACGCTGAGCCGTGACTGTCTTCCGCC
 AGCAGCCCGGTGGCTTGGCGCGGATTTCTGATAGAGTTCAACAGTACAGCGATTGCGCGTCCCTGGCAGCGCCTGGCCGGTTC
 CGCAGCAGGCGCAGGCTGCGTTGCTGTCCGGGTTTCGAGCGGAACACCGCGGTAGCGGCATGATCGGCGCGCGGGCAGACTGC
 GGTGTGAGTTGCGGGCGCCATCGTGCATCAGGTCTGCATATGACCGGATAGGCATTCTGGTTGACAGCAGCAGCAGCAGTGC
 TCGCGGTGCCCCCTCGGCGATTATGCCCCGAGTGCCTCGCGGTGACTCCGGCTCGGGCGTCCAGCAGCGCAGCATGTCATAGG
 CCGAGCAGTGCCTTTCGAGCAGGATCTTACTGGACACGAATCACTACCTGCGCGTGGCCTCCACCCGTCCCGGGGTAACGCT
 GGGCATGCTGCCAACGTTGAGTTTTCAGGCGCGCACGAAGGTCTCGCTGTACTGGCTGATGCCTCCCGCTTCGCCAGTCTT
 GTTCGAGGCGTTCCGATCACTGGATACAGCCCCGTGTTTCGGCATGTCTGCCATCGCGGAATCTTCGCTTACCAGCAGGTT
 CATCGCACTGCGGTGCGCGCAGCAAGCGGATACCCACGCTTGGGCCATACAGGCTGGCCGTAGCGGTGCGAACTAGGTAGGA
 GAGCCCGCCATTGGCGTTCTCAGGCCAGTGCAGCGGACGCGGCTGCGCCCCGCTGAAGCCCTGATACCCAGTGCAGTACC
 GCTGGAGGCGGTGCTGTTGATCACTCCCGTTTGGCACTGAAGGCCAAGCTCGAAGGGCATCTCAGGTTTCGCGCGCGGCAATTAAT
 CTCATGAGGGAATCGACGGGAACAACAGTGGGGCGTGACGTTGGTTACGGAACAGGTGGGATAGCGCTTGGCGTACGCTC
 TTTGTATAGGCGGATGGCGCCCGGCCAGTTGTAGTGCACGAGTTCGAGTTGAGTGGCGTTATGGCTGGCGCGGATGGGTAGG
 CAGGCCGGGACCCCTTGATGGCGATGTAGCCAGCGGGTGGGTGTAGGCGTACAACCGCAGGCGGTAGTTCGGCGAGTAGTAGCG
 GATGAATCCAGCGGAGCGCGGAACAGTTTCGCGCGTACCGCGCTGAGGTTCTTCGCTTGCAGGTTGAAGCCCTCGCGAGTC
 GATATCGAGCCCGCGAGCAGCGCTCGCGCCAGACATCGGTGAATACTGCCCGGTTTCCAGTGGGTAGGCGCAGCAGCAC
 GTTGGGCGAGGCGGATGGCGCGCGGCCAGTTGTAGTGCACGAGTTCGAGTTGCGCGACTGTGCGCTTCCAGTCCAGTGGCGGATGAG
 ATCGTTCGCGATTGGTGGAGAACATCTCGTACACATCGTCTCTGGAGCACAGCGGAAGAACACGCGCTCTGGGTCTGAGCCGGC
 GCTGTCCCGTATTGGCTGAGCGCGCGCACCGAGCTGGCGATCAAGTACCGTTCGGGCTGGAGCTGGCGTCTGTCAGCTTGAT
 CACGCTCGGTAGACCGGGAACCGTTGCGAGGTATCGCAGGCTCCGGCCAGGTGCGCGCAGAGCCTTACCAGGACGATCTG
 GGTGTTGAGGCTAGTCGTGGCGTTGCTTACCGCGGTGCTTGTAGCGAGTTCGCCGACTGTGCGCTTCCAGTCCAGTGGCGGAG
 CAGAGGCGAGCAGCGCAGGATGTTTTTCATGGTGCCTCCGAGACGGGCGTGCAGGTGCCCTGCAGGCGGATCACCGCTGTCCG
 GGTTGATAGCTTCGCGAGGGGCGAGTGTGAGGTCGTAATCCAGCAGACAGGCGTCCCTGGGCGTTTCGCTCCATTGACGCGCG
 AGGCGGCCCTTGTGCGCTTCGGCGCGGCGATAGACCTGGCCCTTGAACCGACCATGCGATGTTCAACCCCTGGCTGTGAGC
 ACATGACCGGCGCGCGCGGCGTATGGCGCGATCTTGCCTCGGTCTCCAGCAGCTCGGCTCTTCGCCATGCCCTGCGGATCG
 AGGCTGACCGGGTTGTAGCGGTAGGGCGAGAGTGTAGGACACCGCGGTAGCCATTGCGGTTGACGCGCGCGCCCTGACCAACCG
 CGTACGCCAGCTCCGCTGGCGCCCTTGGCTTCGACAGGCGGAAAGTGTGCGCGACGTAGGGGCGGAGGTCAGGCGCGCGCGA
 TGCAGCAGCGCGCGCGCGCGCGGCGGTGTGAGTGTGCGGCTGCTGCTGTTCTCGGATAGCGCGCGGCGGATAGCGCGGCGGCG
 GAGGTCTGTTTCTGCAGACTGCCGTTGAATCGCTGCCGTGGTTGTGCTGTGCGCGCGCGCGGCGCAATCGCGTAGCTCAGGCTG
 CGGTCTTCGTGAGGGTGCCTTGAAGCCGCTCTGGTGTGCTGCTGCGCGCGGCTGTGCGCGGAAACCGCGGAGGCCATCGCGCTG
 AGATTGCGGGCTGGCTGGAGGACCCAGCGGATGGAAGGTGTGAGGTTAAGGTGTTGCTACGTTTCGCTGCGGTGCGCGCAGG
 TTGATTGCGCGCGGAGTGCAGGCTGCTAGTGTGCTGCTGATCGCGGTACAGGTTGGTGTGCTGCGCGGCGGATAGGCGGAGGTTG
 TAGCTGAGCTGGCGCAAGTGTGCTGTAGCCGAAGTGCACCTGGGTGTGCGGCTCTTGGCGTCTGATAGTGTGCTGCTGCTGCTG
 CCGGACAGATACAGGTTGCGGTAGCCCCCAGGCGCTGTTGACAGCAGGTTGAAGTGGTTGCTGCTGCTGCTGCTGCTGCTGCTG
 TTCCAGGAGTCTGTTGCTCATCGCGTGGCGCGCGGAAAGCGCGTCCCGAGGTGCGGATAGCCCTCGGTGGAATAGCGGTAG
 CCAGCCAGGTTGAGAGTGGTCTGGGTGCGGTGGAAGGTGCGGCTGATGTTGAGCAGCAGCGCGGCGGCTGCTTGGCGTGGCG
 TTCTCCACCGTGGCATGGGAAAAGATGCTGTTGAAGCCGAAGGCTCCGTAGGGCGTGGCGAGCAGCCTCCGCGGAGCAGCGCC
 AGATAGTCTTCGCGCAGGCGCGAGCGGAGGTTGGCGGTTAGCGAGTTGGTTCAGGCGCGCTGATAGGTGAAGTTCGCGGAACAGG
 TCGTTGCGGTGCGCATACTGGCGCGCTTGGCCGAGGTGGCGCTGTAGCGCGAGATGCCCGGGCGCATGGATTGGGAAACCGCG
 GAAAAGGCGCACGAAAAGCGCGAGACGCTGCGCTGCGGCTCAATCACTTGGACATCCAGATCCGCGCTGTAGGCGGCGCATAG
 AGGTGCTGTATGACGAAGGACCGGGAGCGAGCGGCGGATTCTGTTGACCTTCTTGGCTGATGACCAACCGTGGCTGTG
 CTGTTCTGCTGTACCGGTACCTGTGGAGCATAGCGCGGTGCGAGTTGCGCAGCATGCGGTCTGCTGCTGCGCAGGCGCACGCGC
 CGATAACCCAGGCTGCCGAACAGATTGCCCTCGGTGAAGTCTTCGCCAGGTCAGTTCTGCTGCGCAGGCGCTGCGCAGGCGGCG
 TGGGCTAGGTGCGGATGCTGTTCCAGTCTGCTGCGGGTGTTCGCCGATAGCTGAGTAGCTGATGTTGAGTGGTGGCGCAGG
 CGCCACAGGCGGAAATTGATGCCCCGCTCAGCCCCAGGTAGCCATAGTCCAGTCCCGTTGCGGCTTACTCGCTCGAAGCTC
 GAGCGATAGAAGCTGGCGTCTGATGACGAAGCCATGCTACTGCCAGCGTCCCACTCGTCCGGACCCAGCTAGCCGCGTGGC
 TTGATATCCAGCAGGCGCTGGGGCACCGAGAGATCGAGGCGCAGAAGGGCGGTGTCGAGATTGAAGGTCGAGCCTGGCAGGCGT
 TCTTCCAGGCGCGGCAATCTCCAGGCGTTGACGCGGCTCAGAGGCGGCGGCGGACGCGCGTCCAGGCGGCGGCGGCGT
 TCGCGGAAGCAGGTTTTCAGCGCGCGCGCGCGGGGCGGAACTCGATCCTGGTGCAGTGGCGTAGCTGCCGTTGAGATAGACA
 TCGACATGGTAGGTTCCGGGGTCCACCTGGTTGGCCCGGTTGAAGCGTTCTAGGTTCCCGCGCGCGGAGGCGGAGCCCACTAAC
 AGGCTGTCTGTCGAACGATACTGCGCCTCTCCATTGGCCAGGCGCGGAAGGCCAGCGCTAGGCGCGGAGCCAGCGGCGAGT
 TTGCCGCTGCGCGGCTGCTCAACGTGGCGAGAATGTCGACCTCAATTCGCGCGCGGCTGATGATTGATCAAGAGGCGGCTTCAA
 CTGGACCGGCTCTGGGGCGAGCGGCGAAGGGCGTTCCGCGCGGCGGCTCGGCTGGCCTTGGGCGGGACCATGCTCGACAGCAG
 CCGCCAGCGACGTTTCGCGGATGCTCAGCATCGCGCGCGGCGAAGGAGCGTAGTAGCCGCTGGGTTGTCCACCCGACACGCCA
 GTTGCCGCTCGCTGCACCAGGCTGAAGTGAACCTGCTCGACTAGCTGCTCCGAGCCGCGAAGCAGCGCGGCGGGGCGATAGAA

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND
 PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September 12, 2003 Serial No.: Not Yet Assigned

Page 6 of 118

Customer No.: 21559

FIG. 2F

CGGATCCAGGGA CTG CAGGGGCTGGAGAAACGATGTCGTTCCAGCGGATGTTGACCTGGCCGAGGCCGGAATCGACGCTCTT
GGCCCCGTAGAGGTTGATCGCCGCGAGTAGCGTGTG CAGGCGGCGGTGCGGGTCGCGTAGTAGTAAGATTTCCCGGCGACGTT
CAATGTCCAGGGCCAGGGAACGTATTCGCCGCGCAGCAGGACCTTGCTCTCGGT CAGCGCGACCGAGTAGAGCACTCGGCCGG
AACCCTTTGGGCAGGGCGATTGCTTGACGCCGCGGTGGATCCACGGCGGCCCTGAGCGATGTTGACCAGGACAGCAGGGT
GGCCAACTGGAGGGCTCGAACTACAGACGTTGCCATCGGCCGTTACCTCTGTAC CAGGGCGGAAGGGCTCCGCCAGTCTCT
AGGGCCATCCAGCGGCCCTCGTCATGATTGAGCGTGATCTGCTTGCTGCGAACCTTCTTCGGGTGATGCCGCGGAGGATTGCC
CAGCGCCGCACTCGCTGTCGTCGTTCTGGCTACCGACGAAGTAGAGGTGGAACCTCTTTCTGTCATGTTG CAGGTCGCGGACC
CGCTGGATG CAGGCGGTG CAGTTGTCCTGGACGAACAGGGCCAGGCGCCGCTGCTCTGCAACGCTGGGCTCATGTTGCGCGTG
CCCGACGGGTGCTGTGCTGCTTTCGGTGAGGCGGATCACCCCTCGCCTGGATAGGCGCGGCGGAAGGCTTCGTCGATGCG
CGCTGGTAGGCGAGTTCTTCTCGACCCGGCGCGTTCGGCTGGACCTGTAGATCGGCATACCGCGCGGCTTCTCTGCCGAT
CGCGCCTCGATGCCAGCGCGGT CAGCGGATCAATAC CAGGCGAGTAAGCGCCCGGGGGCTTG CATGAGCGTCTGGTAGCGG
GTCCATTCTGTTGGTT CAGTCCCACTCTTGGGCCAGGTGCTCTGAGATGTCGTTGTTT CAGCGTGCCGAGTTGGCTACCG
GACAGGTCGCTGCTGCTCTCGGCTTGCCGCGCCTCTGGCGGCTGCGGGATAGCTCAGCAGCATCAGGCCGGA CAGG
AGCGATACGCTTCTGATCATGGGCGGGTCTTATGTTGGATGCGAACGCTGCGCGAGGTGCCGCGACGTCGAAGTGCGCGGT
ACCATCGTCAAGGTGCTCAGTCGCCAGGTGCTGCCGCGCAGGCGATCTCCCGCGGAATGAGGTAGATCTGGCTGAGCTGGGT
GGATCCCGGAGGTGCAACCGACAGAAACCGTTCTCCCGCGGATACTCCACGCCAAGGATCGAGAAGGGCGGGGCTCCATCGG
TTTGGGTTTCGGCTTGGGCTTGGCGGGTATGGGGCTTTTGGCGCTGGCAGCATCAGCGCGGCGCCTTGGCTGCTTTTCTTG
AAGCGTGCGGACAGAACCGTCCAGTGCTCTCACGGTTGCTT GAGCA CCAAGAGGTGCGCGGCGGAGGCGGTGCTCTGAGCCAG
GTTCTCGACGGCATCGGAGGCTGCTTGGCGAACGCTGCGCAGCGTCAATTGCTTGGACAGCGCTGCTGGCCTGAACGGAA
GTCTCTGTTGCTGACAGGTGCTGTCATCCACGGCGTCGAGGCGCTGCTGCACTCGACTCAAGCGAGCCAGGATCGGCTCCAG
CGAGGCTTTTCCGGCGCACTGTCCA CGCTGATCGGAGTGAACGAGTTGGTACTGCTGGTAGCTCAGCAGGCCGCGCATTTGC
CGTCAACAGGAGGCCAAGAATCATTGATG CAGGGATGGGCGTTTCACTGAGGGTCTCGCGCGGAAGTCAAGCGCGGCGATGCTCAG
ACGGACAACGTTGATG CAGTAGGAAATACTCGGCGGAAGGGGGAGGATTTTGGAGGGGAGAAATCTTGGCAATCCTCCGGC
CTAATGCTCTGTGCAAGCTGACGAAGCTCGGGGGCTCTTCTGTCGCGCATCAGGCTGAACGCTGGATAGCACAGCAGCTGAC
CTGCACCGAAAAACCAAGTGATTGGCATCGCCGCTGCTGCGGGATCGACGATGGGGCGGACTCGTTGAGGTGGAGTAGTCA
GTGAGGTCACTCGGCTTACCGGCGAGTGCAAA TAGGATATCCCGCTCGGTGACCGAGAGGATCAAGCGCGGCGCATGCTCAGA
TGCTGTGTCGGGTGTGGTGGCAGAGAGTTTCCCGGCTGATATTGCGAGGTGGGGAA CAATAGAACCAAAACCAACGTAAGCAG
AAATGAGTGCTGCACCAACAACCGCGCCACAATGCCACTATCGATTGACGATACCGGTATACGTAAAAGCATGGAGAATCAG
TAGTTACTCACCCAGACGAAGGTTAGGTTGACCATTCGCTGCTTCTGAAGTTATGGGCGATAGATTATGGCCGAAGCTAT
CAGAAAGGATGAGGCCAAGAATCATTGATG CAGGGATGGGCGTTTCACTGAGGGTCTCGCGCGGAGGCTCAATCGATACAAAGTCTTT
ATCTGTG CAGGCAATCTCCAACTGTCCCAACTGAGGTTTCTTTCGGAACCGTATACGCTATGTCATCGTACATAAAATGC
ATTAAGTATCGGTCAACGCATAGATCCTGCAATCACCCAGAGCTTAACGAGTGCCATATCCAACTGGCAGGTGCGCCGATTGC
AGTAAGCGACATTTACCAAAAAATTCATGA AACCACTGAGAACACCTGTTGAGATGGGCGTTGCTCCTAATAGCATCACCTT
TGAGAGTATCAGGCAACCATAAAATCAGCAAGCCTCAACATGGTTCAAGATATG CAGGATCAAGCGCGGCGAGTGTGGA
GGCCCTCCAGGCCAATATG CAGTTCTGTATGGA CAGGAGATAAATACTGATTTTCACTCGCTCGTAATGAACCTGCTGCTGGGCA
GAGAGCGAAAAACCGTCGCAATAGTT CAGGGGCGATATCACCATCGGGTACGGCTTCGATACCTTCGTG CATAAGCGTCCGAGCT
AAACTCTTTGAATCTGTTGTTCTACGCGACAGAGTATTAACCTGCAATTG CAGCTATCAACGTCCGACCCAGGCTTCTGGAG
CGTCTATGCTTGTGTTGGGACAAAGTCTCAGGATGACGAGTGAAGTGGGCTTACTCTTTAGTG CCAAGCGCGGCGAGTGTGTAACG
CATAGCAAGCAACAGTTTGCAGGTAAAGTGAATGGGCTACCCCGAGCTATCAAAACGGTTGCGCTTGATCTATATTATCAATA
TGGGCAGACTGGTAAATTTTCAAAATTTCAACAAGCTATAAATAGCCATGATTGGCCGGCAGTCATCCATGAACCTTAGAACTG
GAATGGTGTACCGAATGATCCTCTCCAGTTCAATCAAAACGATGGAAGAGCGAGCCAAAGTATCTGGCAATATCTTCAACTA
TGACAAATGAGGATATGGAATAGAACACACAGTGAGCGAAACGCAACAGATCAATATTTACCAAAATCCGGGCGAGTCTATTT
CCGGTCTCTACAAGGGGCTGGCTAACAGTGCTCTCTGGCCAGCCATTTCCAGAGGTACAGCTTGTTGGAGGCTTGGGATATCC
CTCTCGTACTCCATCCGAGTTTGTGCTAACGAGATGTCCTGAAAATCGATAAGGAGTACGGAAAGCATCCTTGCTGCTGAGT
CAGCTCAGGTTATCTGCTTCACTCCAAATGGCTCAAGACAAGGCTAAGGCGTGCGGGGAGGTTACAGCCTTGATCAGTTCTG
TCTCTCCAATCTCAATACCATTAAGAGTGCATGGTGCTAATTATCTAAACCTGCTGAAACAATCACCGAACCGATACCCGA
CTAGCGTCGGAGTTGAGATCATGT CAGGTGGCAGTCCGAACAGGATTCTGGAATCGAGGTCTCTACGGTGCCAGTCTCGGCC
GTCTAACTCAATCAAACTTCAGGCGATGAACTGCTGCGCAGTCTCAAA CAGTTGCTCACTCAGGGAATCGGTGTGAAGCTTT
CTCAGCCTGAATATTGGCCTGCTTACAACAACATAGCCACTGGTATTCGTTATACAAACCGAGTGCGGATAACGTTGGCCTATT
GGGCCACGGTTTAGAGATGGTTGAACATTAGGGAGAAAAAGGATAGGTGCCAGATCAATCTGCTCGTTCAATTTTGTGTCATA
TTCCGCGCCCACAATGACCCAAGCTGCGAAAAATACCAGCAAAATGAGTACTCATTGGGGGATGGAAGAGGCTACATCAATATCTGG
CCGAAAAGGATGAGGCTCAGGCAATTTCTTATCCATAATGATGGGCTAATGGGGCTACATGCAGCCTTAAAGGCACCTCTAGA
GATAATAAAGGAGTGGTGCAATTCGCCGATTCCTCTGCTTCATGTTTGCTAAGTATCACCAGACAGGCGTGTGTCAGTAAGC
GTCAAACGTGAGGAAAAATTCGCCAAGCTGCTGTCATGGTGCGGTCCTAGAGTTTGGTTTGAAGGAGCCTATAGCGCTCCCGCC
AAGGGCTGCTACTATATGCAAAATAAGGAAAAAACTCGACAAATGTTGGGTATGATTGAGAAAAAAGAGCTTGATGCCGCTCGC
GCCTTATCAAATAAGCTTTTGTG CAGACTGCGCAACCGAGCTAGCCTATCCTGCCAAGATATACTTGACGAAACACACTTGCCATG
ATCAGTGCTGAAAAGGAGAGAAATGCTCGCTGTTTGGAGTATGCCATCGGGTGCAAAAGCAAATCTCTGTAAGAGATGACGGC
CAACCGCTGGAAGACTTGCTCCCGCGGAGCAGCTTTGCTGATGGAAACACGCGCAAGGCTGATGCTGTCTGAGCGATGCG
AGCGACGAGAAATAATGACGCGTCGTAGTAGGCGCTAGTTCACTAGGAGCGATCTTTCCGACTCTCCCGCGGCGATGAGTCAC
TCGTTCTGTGCGGAAGTGCTGGTAGAGCGTTTCCGACTGATGTTGAATTGCGGGCGAGCTGAGCCTTGGGCTCGCCGCGCGCT
GCCCGCAGCCCCAGGCTAGCACCTGCTTGTGCGGAAGGCGCTCTTTCGGGCCCCGGTACGCGCAGAGCTTGATATCTCGCGCT
ACCGCGCTGGTGCTGCTTGTATGCTGCTCAGCAAGCAGCGCGGCTTTTCAGGGTAAGCGTACGCGCAATAACCTTTACT
TAGGTTGAGCCGTATCAGCAAAAGAAATCGATTCAATGTGTGAGGTACCGCCTCATGCACATTGGCTTTGGTACCGACATCTAC
TCGGAGACCTCTGTCCTTGGTGAAACCAATGTGCAACCATCTTCGTAGGTTTCCAAACAGAGCTACGCATTGATAACGAC
TCATCGCCATTAGCATTTTGATGACGGCATCGGTACGTTGCTAACCGCACACCATCTGCAAGCACACGAGTTTGAAATCCTCG

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND
PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September, 12, 2003 Serial No.: Not Yet Assigned

Page 8 of 118

Customer No.: 21559

10/118
FIG. 2H

CGCCCCATGGCCTGGCCCTGGGCCGAGCCGCTGTTCTTGTAATCATGTCCCTCGTCCAGCACCAGCAGGTGCAAGTAGCCATCAG
GTAGGTAGCGCTTGATGAACCTCGTCCGTTGGTAGCCGCTTACCGAAGCCGAACCTCGATGTTTGCCATCGATCGCTCCATGC
GTTTGGCTGCGCATCGCTGAAGACGAAGTTGCCCTTGGCGTCCATCAGGTTGATGAACTCCGAGACGTTGTCCACCAACATCG
TGGCCAGGAAGTCTCGCCGAAGTCGTTACGAGGCGCTCCGCCCTGACCGGGCCGATGGTTGGTATCCGGCACATCGACTTGA
GAATCGTTGCGCGCCGCTTGGCCCGCTCGGCTTGCCTGGTCGATCAGCGTCCAGAGCGCCACCGGCAGGAGGAACAGGTAC
GTCGACGGTCACCACGCTCGAACTCCTCCACCGTGACCAAGTTGCTTCCAGGTCCTCGAGGACCTGTCCGCAATCCGGGCACG
CAGCGAGCAGTTGGCCGCGCGCGCGCTTTCTTCCAGCAGGCGAGCCGCCAGTGGAACCCATCCGCATCCGCACGCGGCCGA
GGATGAAGAACTCCTGGCGCCCGTCGTAGGCGTCGCCATCTGATCTCGCAGCTTGAGCAGCTTGAGTAGAGTATCTGGGCCAT
TGAGTACCCAGACGCGGGCGGTGGGATGTTCTCCAGGATCTCGCGGCGCCACTTGTAGACCAAGTGCGGCGGAGACGACCA
GGGTCCGGCGATAGCCGGCGCGTGCATGACCGCTGCGACAGCATGGCCATCATGGTTTTCCCGGTGCCCATCTCGCGGCTTGA
TGATTCGGGCTGCTCGTTACGTTCCAGCAGCAGGGCGGTGATGGCTTGGACGACCTCGGCCCTGGGCCGGAAGGGCTTGGCT
TGAGTCGGTCCATCACCACCTGGCGGTGAGCGTTGACGAGCCGCTATAGACCGGGGGTGGAGCGATTGAGCGACTCCAGGA
GTGGCTCAGCGCTCGGTCATTACGACCCGCGCAGGGCTGATCAGCAGCGGTCGGGTGCGGGTGCAGGAGTCCAGGTTGGGCTGA
TCATGTGCATCTCCTGCAGGTGAAGAACGAGCAGGGCGTTCACGGTCCCTTTGGGTATCCGCATACCCTCAGGGTGAGGTG
GAAGGAATCGCCGCGGGGCGAGGTGAGTAATCGCTAGGAGCAAAAGCGTGGTTGAGCTGCGGTCTGCCTCAGTGATGATCCA
GAGCCTGGATTGCTCGCCGCGCTCGATGTCTAGGACGACAGCAGCGCATCGCCGAACCTCAGGGCTCTCTGGTTGGTGTCCA
GTCTTCTGTTCCAGCGTTCCCAATCTCCCGCCAGATGGCGCTTTCAGCAAAGCGCTGGGTTGAGTTGACCGGTTTCCACCAG
GTGGCTCAGCGCTCGGTCATTACGACCCGCGCAGGGCTGATCAGCAGCGGTCGGGTGCGGGTGCAGGAGTCCAGGTTGGGCTC
TTCAGTCTCTTCCGTGGTTCGCGCCGAGGAGCTGATGGTCAGCAGCGGCCCTGATTGACCGAGGAGGGTGTCAATTTCCCATGC
CCGGATGATCGGGATGAAACCGTCCGTGAGGATCCTCAGCTCGGTGATGTTGCCGTCGTCGTCCTCGGTGAATTCGGTCTTGGC
GACCTTGTCTTGTAGGTGTCACTTACGACCAAGATCCGCGGCTTTCAGTCCGACGACGCGCATATCGCGCCGGCGCGC
AGGGCCAGGCGAGGTGCGAGCAGCTCGCGGATGCTGGGCGCGGCTGCGAGCCGCTTGCAGGAGTCCAGGTTGGGCTGAA
GTCAGGCCAGAGACCTCGCAGCGCTGGATTACCGCGCAACTGCTCCGGCTCCAGGGTTACTCGGTAGAAGTGTCTCAGCTC
GCTGGTGGCCGCGCAACACGCTAGGGTTCCACGCGCAAGCCGCTGGAATTTCTCGGCCTTTCTTGGCCGCTCCGATCGC
CTCAGGCGAGACCTCAGTGTGGCTCCGCGCCGCGGCGAGGTCCTGCGACGGAACCGGATGCGGAAGATCACCACTGCTT
GAAGGTAGGATCCGCGGCTGCGTAGATGCGCAGGCGCGGTGAGTGGTGTGCTCAACACGACGCTCATCTGCTCAAGACGTA
GTGAGGAACAATCAGAACCATGACGCGCGCTACTGACGCAACCGGACGCGCTGGTAGAACGCTTTCTCCAGACGCGCGCG
GCCGCTGCCCTGGTACTGCGACGACCGGAGTGGTCCGCCACAGGTCGCCATAAGGCGGGTTGAGCCAGAGCAGTCCGAACGA
GTGCTGCTGATCATGGTGTGAAAAGGTCACTGTGACGACCTCGGTCAAGCAATCCTCGGCATGGTCCGCGCTCGCGGT
GTACTGACAGCAGCGGCTTGGACCTGATCGCGCCGAGGCTGCTGCTGCTGCTGAGTGGGTGAGAACCTCCATCACCGGCGAGG
GTCACAGATCCTCATCTTCCGACGCGGCAAGAGTGAGGGCTGACAGCGCGTTCGAGGGTGACCTCATCGGTAGGGAAGTA
GCCGTTGCGGTGAAAGTTGCGCGCCAAGCGCGGGAACATGAGGGCCATGGGGCTCCTTGAGTCTGAGAGGATGGCTCAGGCT
TGCGCTGAGCATCGGTGGTAAGAATGCTTTCCGATTACCTCAGCGAGGGTGGGCTCAAGCAGCTGACCCCGCAGGGCGAGT
GCGCGCGCAGCTGCGTCCGATCGTCCCGGTAGGCGCTCAACATCTGGTGTGGGTGAGAACCTCCATCACCGGCTGCGCG
CAGTGTGACAGGAGGGGAGCGGACAGGTTCCATCACCAGCGGCCAGAGCCTCTGGTGTGGATCCTCATCGCGCTGCAGGAGT
GCGAAGCGCAGGTTGATTCCTCGTCCGGCGCCGAGGCGCGCGATCGAAACAGCCAAAGATTGAGCAGGCTGCCGAACACGTC
CCGCGAATGACGGGTGGTGGTCTTCTCCAGGAGTCTGGTTGGGAAGACAGGTAAGCGCGCGCGTCCACGATGATGTGG
AAATGCTCATGATTTTCTTCCCGGCGAGCTGAGCTGCGGCAACTCTGTGTACGGTGTGCGCGCCGCGGCGGAA
AGAAAGACAGGTTGCACTGCTCGTCGACACGCGAGGCTCGACGTAAGGTCGGAACCTCTCGATCTGGTAGAGCGGGGTG
GGACTGGGCATAGGAACCTCCTGGAAGGAGGAGCCACGCGCCCTCAAGGGGCGGTGAAGCCCTCGGGGTGTAGTCCAACTGG
TGCAGGGGAGCGTGGTGGTACCGAACGCTCTTGGTGTGAGGTCGAAGCTGAGTCCATCGGCTTGGTGCAGCGGCTTGGCC
GACAGCGGTAGTAGAACCGGCGCGCTTCCGGAATGGTCTTGTAGCCCTATGCGCGCGGCGAGCTCCAGCAAGTGCAGCT
CGATCATGGTGATCGAGTGACCGCGCGCTTCCGGAATAGAATCCGGCAAGGTGAAAAGTTCGTTTTACGACAGTCCAAAGTT
CATTGCCGAGACTGCGGTGTGCCAGCACCTCGGTGCTCCTGCGAAGGTAAGTGTGGGGCCAGCAGCTCAGCAGCAGGTCTT
CCTTCGTCTGATGTGAGAAAGCCATCCCATAGGTGCTCCTGTGGTTGGCCAGGTCCTCCTGACCGGAGAGCCCGGCGAGGG
GTGGTAGGGTGTGCTTACGCTGTGTCGCGCTCAAGGCTGGAGGAGCCGCTAAGCTGTTTGCCCTGGCTGCGAGTTCGCGAGT
ACAACAGGGCTCGTTGTGCGCGCATGCTCCGGCGAAGGGGTCTTACCCGCTCAGCAGTGCGATGATGGCTGGTTCGACGA
CGGATCGGTGAAACACTGCCGCGCGGAACGTGTCAGCCAGATTGTCCAGTCCGACCTCGCGCAGCGAGGCGCAACACGCAACA
ACTCCTCGCTGAAGCCTGGGTGTGTAGATCGATTTATGCGGCCTCCTTGGGGTCCACACCTGTTGTTTGAAGTGAAGCTC
ATAGCCAGCGCGCGAGTGCACGCACTGTTGGCGTAGTCGCTTGGGTGCGAGTCCAGCTTGAAGTTCGCTAG
CGGCCAGACCGTCCCGAACAGTTCTGCATCCGCGTCTGTGTCCAGAGAGGCGGCTGCTCTGCAGGAGCCGAGTGTCCATACC
GAAAGCGCGTCCCTGGTTGACGCGAGGCAAGAGGCTTCGACTTCTGGGGTGACGGCGACTTGGTGGTATCGCCGGTACTAC
TGTGGGAGCTGGATCTCGGCACTTCTGTCAGCGGATCCACCTCATTGCGGAAAGACGCTCTACCTCTCGTCTGCTGAGGCT
TGCCATATCGTTACGCTCATGCTGTCCACTATGGCGCGGCTCAGTACGAGTACCGGACCGCGAGCAGCGAGGCGGCGG
GATATTGGCGATGACGAAATCACCGCGGTACTTGCCTTGGTGTGTTGATCCAGGAAGGCATCCTTGATGATGAATCCCCGAT
CGAAGTCCAGAGTTTCCACATTGAACCTCGCCATAGCGACCGTTGATGGTTGCAATGGCCAACTGTCCGGGAAGTGTGATCAT
GGTATGACCTGCGAAAAGGATGATCCCGCGGAAGCGGATCGTGGTGAGGGGATGAATACTCGGTAACCGGGTGTTCGGCGAG
GCACCTTCTGCAACTCGATGCGCAGGTAGGAAGGCTCATCGGACGGCTCGACTATCGAGAGAACCGAGCAGCGCAGGCGCG
GTCCTGCCCTCTTCTCTTGTGGTACAGGTGGAATTGATGTGCTTCGCGGTGCGCTGGAGAGAAAGCTCCTGCCAGGCGGT
TTGAACGACACTGGCCAGCGGGTGC CATCTCGTCGACCTGGCAGGGTGGCGAGATAGACCAGGCGCAGCCAGGCGCGCTT
GGTGAGAACGATCGGGAGTTTGGGTACGCGCAGTGTGACCTGCGCTGGAGTGGTTGTTGAGAGAGGTGATGAGAGGTCTCCG

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND
PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September 12, 2003 Serial No.: Not Yet Assigned

Page 10 of 118 Customer No.: 21559

11/118
FIG. 21

AATACGCAGGCCTCTCCGCGGGGATGAAGTCCCCGCGTGGGTGGGTAGGATGCACTCCATCACTTGCCTGATGCGACCGCGCT
GCCGAGTGCTGTGGCGGTCTTGGGTGGGTAGCGCTTGGCGCGCGGAGCCTTAAACCTGGCTCCTTGGGAGGCGCTGTCAACG
ACAGCGAGCCAATTGCAATTCGGGGGCTGGCGGGATTTCGCGCGCGCGGATTGCTCAATCGTCATAAAGCGGCAGGCCCTGAG
CGTGGGCGCGAATGGATCAAGTACCAGGAGCCGAATGTCTGCTTGTCCCGCGAGGTGAAGCAGATCGACAAGGTGTTTCGTCGAG
GCCGAAGGCAAGCTGTTCTTGTGCAAGAGCGGCTGCGGTATCTCCAGGATCTCCAGACACTCATCCGTCCAAGGGGTGTTGTG
CAAGCGAAAGCCACGGCACCGTGCTCGTTACACCAGAAGAGCGATACCAACACCGAGGGCAAGCGGATTGGCATCTTGGTCTGA
GAACCTTTCGTAGAGGCTATTCCACGCGCTGATGGGTAGATGTGTTTTCTGATCTCCAGCAGGGGTTATAACCGCGCGCGAA
CGAAAGGTATCGAACGACACTCTCGGCTTCGGTCAGTGAGTATTGATCTCCGATAAGCTTTTTTCTGCGCTCCTGCTCACCGTA
GATGCTGTAAAGTACGTTTCTCACTATTCTTGGTGAGGCAATTTTCAACTCTTTCATCGTTAACTCTTGGATACACCGGC
GAAATCGTTATTGCAAGTAGCACAGCTCGACTGAAAATATTCGTCTGGGAGGTGCTGCTGGGTTTCATGCGAGCGCGGTAGCA
CGGTAGATTGTTGCCTTCGTAGTGATCTGCAGGTTTCGTCTGATGCGGAACTCTCATAGCCGCGAATAAACGGGTGAGTTG
AGTCATGGGATATCTCCGGTTGCGGGAAGCCCATGCCCTTGGAGGCTATGGACCTCCAGGTGAGTAAGGGGCGCCGAGCGCC
CAGGTGATGTCAGTTGTTGCACTTAACCGCATAGGTGTTCTTTCAGCCGAGCCTTCTGGTATCGCTCTACAGGCCATAG
TATTGAAGTACCGGTGATGTGCTTGAGTCGCAACCATGTCACAGTCGGCTGCGCCAAAGCTTCGATTGCTATCTTGAATTACA
TAGAGCTCGATACTGAGTCCGGCATCGCCAAATTCATCAACCCGGCTTTGCACTGGATGTACTCGAACCTGCTTTGTGCGACA
TGCACTCCAATGTTGAGTAGAGCTGGGAGCCTTTCTGGATATGTGTTTCATGGTGAGTTCCAATTCGTAATCAAAATCGGAGAGGT
CACCTCCAGCAGGGGTGATGACCCGCTGGGTGAAATAGAAGTAGCATGGGCAAAAGGGTAAAGCTTAATATTCAGTT
CGCTCCAGTGGCGGTGGTCTAATGCCTGTTGGCAAGTTTCGTCTGAGGGGTAATATTCAACCGCATCGCGAGCGCGGTCTCT
CTTCGTGAGAAGTCCGATGTAGTAAGCACGGTTCCTTATCATAAGTACCTGCAACGGCAGTTCAGGCTGAGATGGGCGAGCG
CCAGCTTTCGCTATTCGATGGTGATTCTCCTTTCGGTAGGAAACACAGTCCCATGGGGACTGAGGTCCCTAAATGGGAAATG
TTTGGCGATTATGCGAGGAGATACCTGGGCATTGTTGTTGGCGAGAAACAGCCAGGCGAAGTTGAGAACTGACTGCTGAGC
AGCAGTCGTCTCGACGTCGCTGAGGCTAACCTCTCAACTCCGCGAGGGCGGGCAGTTTACGATTTGGATTGTTGCTGATG
ATGATCTCCAGATAATGGAGAACACACGCGCTGGGGGAGATGGTGCCCCCGGATGGGTATGCGATGAAGAGTGGCGCATAG
CGTAGGTAGTGAAAGGCCCTCGGATGAGGGCGGTTATCAATCAAGCCGTGTTCAAGCAGCGGCTTGTTCAGCGTGTCTAGCG
GGGCTGCGGGCTCACCTTGAGGTTGCGCTTTCGCTTGTGCGCTTTTTCAGCTTCTTCTTCGCTAGACGGTGGTGCCA
TCCACTTTGATCCAGAGATGAAGAGCAGCGGCTTTGAGGCTTTGAGGCTTTCGCGAGGGCGGGCAGTTTACGATTTGGATTG
AAGGGATCCGCCCAGATATCGCCGATACGGAAGGAAATCAGAACCTTCTTCTTGGCCTCGACTGCTTCTTGCAACGGCGGACA
AGCTTTTCAGCCTGGGCGCGGACCACTTTGCACTGATGATGAGGATTTCCACGCTGTCTGCGCGCCATGGAGGGCTGCGAGC
GTTACGGCGAGGAATGGTTACCTCGGCGGATCGGTACCTCGCGGATGCGATTGAGGTAGCCGATACCGGTGGTGTGCAAGTCA
AAATACTTGGCTTCTGGGCTTGGGTGTTGTTGCTCATGATGATGAGGATTTCCAGTTTCGTAAGCGAAAGCGGAGAAACACTTGGCC
GTCAGGAAAAGTATTCCCCGATGGGTGCTGTTGGGAGGTAATCAGAGGGAGAGGGCTTCCATCAACCTGTGGGTGATTG
ATCGCGCTACCGGAGCTGGTGCATCTTGGGTTGATCTACAGGAAAGTCCGTGTCGTAGTCTTTGAGTATGACTACTGGG
GGAGGTTGCGGACAAAGCTCAGCAGCTCACGTGGTAAGGATGGGCAAGGAAAAATGAATGTCAATACCCATGAAATGAAAAA
AGCGCCAGGTTGGGCTTGGGCTTGGGTGTTGTTGCTCATGATGATGAGGGAAGAAATCCGATGTTTGGGCAACAGGTGAGTCA
ACTGAACAGAGGGCTATGCCGAGCAGATAACCAACAGGCTGGTCATGATGATCATGCCGAACTCGGTCTGGGACATTCTACT
CCTCCTCGGTGAGAGCGATACCTATGAGCCAGATTACCATACCCATGACGAAAGCATATGCCGTGAGCAAGCGGTGCTCGCTGA
GGAAGATGCCCAAGGCTCGCCGCAACGGCAGTGGTACCGATTTTTCGGAGATCGTCTGCGAATGCAGAACGTTTCGATTCA
TGGTTTCTCGAAGAAATGCCGGAAGAACCCCTCGCGGATGGTTGGGCTTTTTCGCGAGGGTGGGGGTGGTGGTGGTGGTGG
CCGAGAGCAGCTACTTGTTCATCCGTGATGGTGACAAGTAGTCTGCCAGCGTTGTTGGTATTGCTCCTGAGTACGCCGATGG
AGTTCCCAAGAGACTGCGCTACAGGAGGTGCTGGCAAGAGCAATCTCATGATCGTTACAGGACTCTAAGTGGAGAGGGGGC
CTCCCCATTTCGCTTTCGGAGTTGGGTGCAATCAATGGTTCGAGGGAACCATCTTGAAGCTCCTTCTGACTGGCGTTTCAGCT
TTGGCGGTGCGCTTGGCGCGCTTGGGTTTGGCTGCGCTGCGGCTTTTGGCTTTCGACCTCGCGCGTATTGGCTAGCATACAGTCA
AGGTTGACCTTCTTACCGCGGACGATGGCTGGGATGTGCTGATGTGTTCTAGCTCGCCAGCGGTGCGGCGGCGGAGGAAAGCA
GGATCGATGCGCTGACTCTCAACGCAACCGGCGCTTGTGCTTGCACTACGACCGAGGGGCGACTGAGCGCTCTGCGGAACTCA
CGCACAACTGCGGAGGGCAAAGCCGATTTCAAGCTGGTGGATTGCTTGGACATGGTGTTCCTCTGAGGTGAGGGTGGC
GAGAAAACTAGGCGCTGGCGGGAAGTGGATTCCCGCTGGGTTGCGGAAGGAGCTTCTATCGAACGAGCCGACCGACCGC
GCTACCAGAAGCTATGCGGTCTTGGGTTGGATCAACACGGTGGGCTGTGCTGCAGCATTTCCGATCCGAGCTGCATGGGGCAT
GTCGCTGACAACGTGAGCAGATCACGCTTCATCATCGAAGCGGGGGCGAAAGTGTCAAGGTCAAGCAGGAGGGGGAGCGGAT
CGAGCGGGCGGAGCGAGGGCGGAAGAGCGCGCTGGTAGAACCTGTCAAAGGTTTCCCAAGCGGTGCTGGAAGGGGAGT
CAAGTGACCAACCTGTGTTGATTCGTCATCCCTGGGATGATCGTCTCGATCCCATCAACCGGCGATTATCGTGGCTTGGC
CAGCAACAGTAGGCACGATCCGCTACAAAGGGAGCCCTGCGTATCACCGGCGGGCCACCGGCGAGGAGAAGGATCCCGATG
GGTAGCCTGATAGTCTGAGCAGGAGCATCAGGCTACCATGGAGAGGGGAAAGGAGGGGCGTAACACAGTACGACCTT
AAATCGAGGAAACACCGAACCTCTTGAGAACCCTATGTGCTTTAACTCGCTCGACACGAGTTGACGGCTGCA TGACAAGCAA
AAGAACAGGGAGACACAGGGAGAATTCTCAACAGATCGCAGCGCCCAACCGGTGTTGTAGCGTGCTTACGGTTTGGCGATGA
GGGCAACCTGCATCACCGGGCGCTACTGATGACCGTCAGTAGCCAGCGGGGTGAAATCAAAGTTGATGGCATTATGTGTTGAG
AAGGCCGTTTATCAAGCTGCGCGGGGTTTACTGGCTGTGGAGGTTATTTGACCGCGACCAACCGCGCCCTCGGGGATGGGCG
TAGGGAGCGCAATGACAGTGGCTCCGAGCGGCGGCTGACTAACTGGCCGTTGCGCTGACGCCAACAGGTTCGGATTCCGTAC
CAGCATCCGCGCGCGCGAGCCCGCGCGGCGAGGAGTGAAGCGGTAGTGGTGCCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGG
GCGCTGGCTTGGGAAACAGGCGCGTAAACCAAGAGCGGAGTCCGCGCGCAACCGGTGAAACCGGGGTTGGCGTGTGCGTACT
CCGCCAGACTGGACCGGAAGAGCAGCAAGCTCCTGGCCAGGGTGTGACCTCGGCCAGTTCCGCGCTCGAGGTGCGGTGGTGGC
TTTGTACTACTAGCCAGGTCCCGGTGATGAGCGCGAGCACCAGGACGATCCACATCAGCGGCATACGGTAGTCTTGTAGTGGT
GCGACTGCTAGTTTACGAGGTAACGGTCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG
ATCTCAGTGATTAGCCGAACGAGTAGTTGGTGCCAAATGTCTGGTTGAAGCCCACTGCGCAGAAATCGGCCCGCGCAGGTGATC
GCGTAGCGGACGTAGTCTGATTTACCATGGCGCTATCTGACCATGCAACGTACACAGTCTTGTATCAGCGCGCGCGGTGGTG
TTCAGGCGCATGGAGGCGCGGTTCTGTCCTTTCCTGTTTGGCAGGACAGCCAGATCCGGTGTGGTATGCTGCGCAGC

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND
PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September 12, 2003 Serial No.: Not Yet Assigned

Page 11 of 118

Customer No.: 21559

12/118

FIG. 2J

CCGCTTCGGCGAGCAGTTGGCTCCTTCGGTGGCCACGCCTTTGAGCTGTAGGTACTCGCCGACTTCCGTCCGGCCCTCGGCGGTG
AGTTTGCCCGCTTTTCATCTCGCCGCGCGGTGTAGACGTTCTTGTTCATCCAGGAGCGCACCCAGGTGCTGTGCCTCATGTACCAG
CCGCGCCCCCATTTCTCGTTGTACCAAGCGGCTGTACACACGGGTCTCTGAACAGCCCTCGGTGTACGTCTCGCCGGCGACGTCGA
CGAGTGGCGGCTTTCCACCGTACCAGTGGCTGTACGTTGCGCGCGTACGATGTCGCGGCTGGTGGTATGTTGCCGTGGCGCTG
ATCGCCCCGGCTGCGGCGATATTGTTGCCCTCCCATGTCCAGCGTGGTATTTCATCCGGTTGAGTTCAGGATGACCCGGGACGGCA
TTGCGGCTAGAGGTACTCGTTGGCCATGGCGCCGCTCCTGGAAGAACAGCGCCGTCGCCAGATGTCCAGCGCGGGAGAGCGCTACC
AAATTGCTTAAGGCCACCTGCCAGGCGCGCCCTGGGCGATACTGGTGTGGTGTTCAGATGTAGCCCCCGGTGGCTCCCATG
AGCTGCGCGATCCGGCGGAGTTCAGAGTTTCGGAAGCTACCTGTCCACCGGTGGTACAGATCAGCGCTTTCAGCTGGTTGGCGCC
GGCTTGGCGGCGAGGACCTGGTATTGCTGGCCGTAGATGTTGGTGTGCGGGAAGCCTGCGGGCAGGTAACGGGTGTTGCGCAGC
ATCGGGACCGGTGATACCCGCGGGGCGGTGGCGCCGGCGCTGGCCAGAACCGTGTGAAGTTGTCCTTCAGTACTTCTCCGT
GCCTTGGCCACCTGTGTTGCTGTGCTGGCGCGTGTCTGCTCCAAGCCGTCAGGCTAGCTCATCAGGACCGATATGCCAAC
CGGGTCGCGATGGCGATACAGACAGAGGCGCATCATCAGTTCGATCAGATGAATCACTGCTGCGCGTACTCTCATGTGTTT
TCTCCATAGGCTTTGCGGGGACTGGCGAGCTTGGTGTCTGTCAGCAGATGAGCCTCCTTTCCGCGGATCTTGTGCGCGATGGA
CGGGCCGTGTCGCCAGTGAAACAGCAGTGGCTGGCGGTCAGGCGGCGAGGAAACCATCGCTCTCAGCGCTTGGCAGATCAC
AGCTTGGATGCTCTCGGCAACACCGTCTGATGGCTTTCTCGCCAGTGGCTGGGCGAGCGTACCAAGCGCTCGATCGCTGTAG
ACAGCTTTTCGCGTGAATGGTGGATCACCAGGCTGCCGTCAGAGGCGCTTGACGGCTCGAGGCGGTGTCCTCGTCTGCG
GATCTGCCAATCAGCACCAGGTGCGCGCCGCGCCGACGCGTGCAGCAGCAGGGCCTCGCTGTATCCGCCTGAGCGCGGTGAGGT
TCTGACCTGGATGCGCGGCGAGCCCATGCTGACCGCTGAGGTTGGTTTCTGCGGCTCTCGACGGCGCAGCCACCCCGCC
CAACTCTGCGAGGCGGGCCAGGAGCAGAGAGGCGGCGGAGCTGGTCTTGCTGTGCGCCATCTCGCCACTCGCCGAGAACAGGACCGAGCC
CTGCAACTGCGGATCATCAGTGGCGGAAACCATCTCGCTCGGATCGGAGCAGTCTCTTGAACCTCCGCAATTGGGCGCTCGACCG
GCGCAGCACGAAGACGCTGCCACTGAAAGCGTCTTCGAGGAGGGTGACGCGAAGCAGAACGCGCTCGACGACGATGGCGAACTC
AGGGTCTCTCGAGCTCTCTCAGCGTTTGGCTGCAGATCTGAGACGAGTTCCTGAGCTGCGACTCCCATCGCGAGGCACTTCGAC
TCGGCGCGCGCGCGGCCCGGCTTTACGCTCGGCGAAGCCTTCCAGCGAGTAGAGGTCAGCAAGTGCAGCTGGATGAT
CGGGTTACATCACGAACATAGAGTCCAGGTAATGCTGTTGCTGTGCTGCTGAGGCTGGGTCGCGGCTGCGGCTGGCTGATC
TCTCCGTGATCGAGGATCCGCTGTTGATTTTGGTCTGTCGAAAGGTGTTCTTCGCGATCTTGGTGGCCAGCGTGATACAGCG
TCTTGGGGCAACTTGTGGTGGTGATCAGAAAGCCCATCGCGTGGACGAGACAGTGACCGATCCGCGGTAGACGTTGTAGACG
ACGCGCGAGGAGACACTGATGTTCTTCGGCACCGCCTTGATTCGATCAGGCTGGGATCAGGTTGGTACCGCTGGAGCGGTAG
CCCGAAGAGGCTCTTACGGCGCTGCGTGGCTGGCGCAATGACGCTGATGTTGCGTTTCTCTGTTGGCTGGACGAACTGAAC
ATTCCAGCCGCTGCGGCCAGGCCAGGCCGACCCCGATGGCGATGATGATCAGCAGATGATCATCTCGATGGAAACGAAACCG
CCCTGCGTCGGAACGGGAAGTGCCTTGGCTAGTGCTCATGGAGCGCTCTTGGTGAGGTTTCAGTGGTTTCATGGATTGGATGAGC
TGCTGTGCTGTGATGCGCGGAGCAGGACCGATCATCAGCGCCCGATGAAGATCGGCGCGAAGTTCTTCCACAGCCCGCG
GCCAGCTGATCTGCTTGAGGCTGGTCTCTCGGACCGGCGCTGAACTTGACACGCGCTCGGAGAAGCTCCCGGTTGGCG
AGGATGCACAGGTACTGGATGGCTTGGCGGTGGGGAAATCGTGACCGCGCTGCAAGGGCAACCCAAAGTTTCGGCCAGG
CCACGCGGTAGCGGGCAGCTTCCAAGCGCTGCTTCAGCCAGGCGGGGAGATCTTGATCATGCTGGCCAGGCTGTGTAAGGG
CGTATGCGCGCGTGTAGCATGACCGCCATGTTTCAGCAGGAAGGTGGTGCCCTGGAGCATGCGCTAGATGGACAGGCGCGCAGC
TCCGGTCCAGCCAGACCCGCGCTTCCAGCGTAGTGGCAGACGTGACCATGACACACCGTGAGGTTGATGACGCGCAC
AGAACGTAGATACCAAGTCTGTGACGAAGCTGGCAATGCGCTTGAGCTGGCGAGCGGGCCGCTCAGGTTGATGGTTCGGAG
AGCTTGGCCAGGCTGGGGACCATGCGATAGGCCACGATGCAAGCAGGAAACCATCATGGCGGACAGCGCGAGGGGTAGAGC
AGCGCTGCCAGATGGTGGCGGATCTTGGCTGGCGCTCGACAGGGGAGACGGCATCGCCATGGCTGAACAGGATTGCGC
CTCATCTCGCGGCTCGATCAACGCTCGCTCTTGGCGGGGAGATAGAGCCCATGCGGCTGGCCAGACGCTTCCGTTGGAC
AGCCCCATCAGCGCTTCGCGACTGCGCATGGCCACCGGATGAAACGGAATGCTGCCCTCATGAGCGAAGATCTTGTGCACTCT
GCCACGACATCTTCAACGGGACCCGCTTTTCGAGCAGGTTGGACATGCTTTCGTAGAACTGACGGCGTTCCTTGGCGCCGAAC
TGCTTGCTGTGATGAAGGCGAACTGCAATGCTCCGAGAGCCCCCATCAGCACGACCTCTTCGAGTAGAAGCTGTGTCGAGCA
GATGCTCATGAACTGAGTGGCCAAATGAAATCTCGACATCTGCGGTCGAACTGCGCTGATGCGCGGATGGCGT
GGGCGTGTGTTGGTGTAGCCCTGCAATGGTCTTGACCCAGTAGTTGCGTGGCTCCGGCTGGGCGCCCTTGGCGAAACACAGCATGA
AGGCGAGGGTGGGCAGAACACCTCGGCGACGATCGAGCGGCGGTTGACCCCGGAGCCACGGCAGGCTTGGCAGCCAGGCCCTT
TGACGTGAACCTGGGAAACATCGTCAAGCGTCGGAACCGTTCGACCAAGTCCGGGCGCGAAGTTGGTCTTGGTGGTCTTGAAGC
GCATTTTGCACTGGGGCAGAGCTTGGCGAGCAGCTCTGGTGTGATCAGGCGGTCAGCAGGCGCGGATCGAACAGCAAGCCGG
GGCTCGAGCGCCAGGTCCTTCAAGCGCTGGACAAATGCGCATGCGCTGTTGGTGTGAGGTTGACCATAGCGCTGCCCGCTCA
TCGCACCACGGAACCGGCGACAGCGGCGAAGAGTTCGCTACTTCGCCGATCATCATGTAGTCCGGATCCAGGCGCATGCCGT
TGGCGATGCCCGCGCCAGGCTTGGCGTTTTCGCTGTGGTTCGGTGGCGCTCGGTAGACCATGGGTTGATGCTTTCGCTCG
CGCAATCGGATTCGGCGGATCTCGATGGTTCAGGATGTGCTGGATCCGCACTGGAGCTTGTGAGGCTTCCAGGCTGAC
CCTTCAAGGTCATCGAATTTCTGACCCGGTGGGCGCGAACAGGATGTTGATGCCGTAGGGCATACGATCATGCGATCGGAACA
GTGCGTTTCTGCTCGGGCAGGTAGCCGAGCTGTCCAGGCTGTGAGGCGGCTGTGCTGCTAGAGCAGTCGAGGATCATCAGGA
ACCCCGCGCACCGGGCGGGTGGCGATCCGGGCATGAACAGGTTGAGCTTCTCGACGAAGTCTGGCTCATCCGCGCGTCTG
GGTCCAGTTGCGGCTTGAACAGTGGCTCGGCCACGTCGCACTGGATGTTGATGTTGATGCTGACAGAGTCTGTCAGCTCTGCT
TGCGGAATGCTCGACGGCTTTCAGCAGGCGCTCGACGCGGAAACGGAATCTTGTGCGGTCGGCGGGGACTCACGAGAAAT
GCACGTCACTGGCGCGCAGCTCAGTGGCCTTACGAATGATCTTGACCAAGCGCTCTCTGCATCTGGGTGGCGCGCTGGCCATCGC
TATCGACAGGCGGCTCCATGGCGACGGCGCGCTTAAAGCTGGTGAATGTCTGAGCGTCCGTTGAGGTTGAGCTTGGATTCGGAAGC
CACGCGCATCGAGACGCTCGATGAAACGACAGAACTGTATGTCGTTGGTGGCTCTCCGATAGATAGAGCGTGGCTGGCAGC
CCAGAGCATGATTTCTGACGGTCTTCGATACCTCCCATTCACACCGTCCGCGGTGAGCAGTTGGGTACCATGGAAGGCT
GCGCAAGCGCGCAATCTGAAGGTTCTCATGGAGGTGCCTACTGAATGAACCGGCTGCGGTACAGCCCCGGGCAGCGCCGGCGG
AACCAGGCGCGCTTGGGCGCTAGAGAGGCTTGGGTGGGCGCAACGCTGGAGAGAGCCACGGGACGCGGCTTGGCTGCTGTC
GGTGAGCACGACCTGTTCCAGAGAGTTGACTCGACGCGGATTTGCCAGGAGCTCCGCGCGCTGGCGGCTCAACCTCGTA
CCGCGCGCGGAAACAGAAACGTGGCAGTCATCTTGGCGCGCGCGGTAGATGGTCCGACCACTGGCGGAAACCGGCTGGTGTAT

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September 12, 2003 Serial No.: Not Yet Assigned

Page 12 of 118

Customer No.: 21559

13/118
FIG. 2K

CGTCGGCGCGCGCGCGCGCACCGGCATGGCGAACGTCGCCCCCTGGGCGCTGACGACCTGGCCGGTTTCGCCTTTGCCTTCCAA
TTGCCGCTGCGCCGTGGCCAGGCGCACCTTGGCCTCGGTGAGGATGGCCTGGGCTGGATCTCCGCAAGTTCGCCAACCGTGCC
AGCGCAGGCTGGCCAGACGCGAGGAGGAAGAGCACGGCCACGGCCATGCCGATCGGCTCAGTTCGCATAGATTTCTCCTGTGA
CAGTCCAGTCCAACCTGGCTGTCTTGGAGCGTGGTTTCGAGGTTGGTGATGCGGACACCGGCAGCGGGCAGCCCCCTGGAAGGTCA
GGTCTGCCGGGAGCGGGTCTGGGCGCTGAAGGTGAATTTCTTCAGGAAGGCAACACCACTTGCTGTTAGTTCGACGCTCGAGCTTCCG
CGCCAGGAGGGGCGGAGTTGTCTCCTGGCTGATCGACAGCTTGGGCTCGACCCCTTGACCGGTACAGGTGGCTCGTCAGCGCCT
GCAGAAGTTCGTCCGCCGCGAGTAGCGGCTCATCACTGCCGATGGCAACCTTCAGATCGACCTTCAGGGCGCGGTTGTGCCGT
TGTGATGACGAAGGCGGGGCGGTCCGGCGAACAGGTGCTGGCTGGCGCTGTGAGTCCGGCTGCTGTGCTGTGCGCGGTACGGT
GGTAGGTGGCGACACGAGACCTGCCGTGCACTTGCTGGATTGCAAGAGCCAGCCCTGGATCGACAGCGACAGTACCCCGTTG
CCTTGCTACAGGCGCGTAGCATGTCTCGAGGTGAGGTGAGCGTCAAGGCTTCTGCAATGACCGAGGTCCAGCGGTGCT
TGGCCTGGGCGTTTCTTCTCGGCCAGCTCGGCGAGGCGCCTGTCTCCTCGAGGAGCGCGGCTTGCTGGCGAGCTCTTCTGTGT
GGGCATTCCATAGGTAGTAGGCGGTAGCGACCCACCTACCAACGAGCAGGGCCACTGCCGTCCACTCCCTGGCGGACA
AGCCGAAGGTGAGTTGCCGAGGCGGTAGTCGCGCGCAGGCGCTTGGCGCGAGCAGTTCTCGATGTGCAAGTCTTACCGG
GAAATCGAACCTTCCGGAACGAAGACCTGTGCGTTTCCAGCAGCAGCGCGCTAGAGAGTCTCTGACCGGTCCGCGGCT
CGTCGAGGGTGGTGACAGATCTGCCCGGGACAATCGCGCCATCGAGCGTGGCGACAGCGCGTAGCGGTCTCTGTCGACTT
TCCAGCAGGCCAGGAAGTGCCTCGAACCTGGCCTGAAAGCGCCGAGGCCAGGAGTACATCCCTTGACTGCGCCTTGCAGT
TCGAAACGAAGCCGCGCTGGATCACCGTCGGTGAATGGCGGATGGCAACGATGTCCAGATGCTTCTTGGCCAGCTTGGCG
TGGCGGTTATGTCGGTGGCTGGAAGCGCGCCAGAACAGCGCGCTGACGAACTTGTGCGGTGGTAGCTCAGCAGTCTCGC
AGAGCTCTGGTCCACGGCTGCCGAGGTGAGGCTTCTCATGGTTTCCAGCCAGCAGCAGCGGGGTGATCAGCAACAGTACCT
CGCGCTTGGTATTGCGGGTCAGCCCGCGCCAGAACGAGAACCCAGCGTCCGCGGTGCCGACCTTGTGGTGTCTCGGTGG
TCTGGTCAAGCCGAGAGTACCAAGTCTCGCGCTCGCAGCAGTACCTTCTGGTCAACAGTGTATGTCTGATGCTCGGGA
ACTGGGCTTTGGAGTCCCCGCTGTGCTTCAAGCTTCCGCGCAGGAGGTGATGTTGATGTTGATGTTGATGTTGATGTTGAT
CGCTTTCCATCACGAACGGCAGCAGGCTCATGTTGTAGCGCTGGTACGGCGCCCGGATCAGCGAGGTGGTACTGCCGACCT
GGGCGACGTTGGAGATCTGGCTGAGGCGCAGGTAGCTGTGCTAGCGCGCGATCTGGATCGCGCGCGACTGGAGGTTGAGCGTGG
TCACGGACGGGGATCGGACGGTCGAGACCGCGCCCTGCTGGGCGAGCGCTGGACCATGGCCTTGGATCTGCCAGGCGCTGT
TGGCGGTATCGCAGGATGCTCACGAGCGCGGAGATCGCGCTTGTGATCGATCGCGCGCATGGTGTCTTCCAGCCGACCTCGC
TGTTGTTGAGCGACTTGTAGACAGGTTCCAGTGCATCCCAAGTTGATCCTTGTGGTTCAGGGCGACCGAGAGCAGTTGACGT
TCAGCAGCAGCTGCTTGGTGTGCTCTCGTTCTCTCGTTGACCAACTGCTGGACACGGTTGAGGACTTCTGGACGGTGGTGA
CGGTGAGGTCCTGGTGGCACGGACAGCGACATGCGTCCCATGCTCGGCGTCAGCATCGAGTTGATGCTGTTCTCGATGTGCG
TGAGGATCGATGCTTTCAGCTCCGAGCTGGTGGTCTGCTTGTGCGGAGTCCGCGCTGATGCGGAGCTGCCATCTCTGCGG
TGGATCCGAGCGCTCCCGCTGATGCGGCGCGCGTCTGTCATACCGGAACGCACGGTGGAGTCCACCGTGTGACGTGCTCGA
AGGCGTACATCGGGAAGGTCCGAGTGTCCAGGTAGTAGAATCGACCTTTTCTCGGTTGGGTTGTAGCGCCAGGACAGCCCG
CTCGGGCGCGATGAGATCGAGGAACCGCTGACTTGGCGTTCATTGATGTTGTAAGAGACCGCGGACCGTAGGACCCGA
AGCTCGACCCCATGCTGCCGCCGCGCACAGCGACATCGCGTGGCGGAGCGAGGAGGAGGAAACAGGAGTCTGCTGCGG
CGCCTTGGATGGGCGCGGGCGTGTGCTCGCGCGCTGCTGAGGTTGACGGCGAAGGCGCGCGGTTTCCAGCGCTCGGGCGTGA
TACTGACCGCATGTGGCATTGGTTGATGACTTCTGGGCGGCTCCTGCGAGCAGCTGCGCGCTGACGGGCGCCACGTACGA
TGCACTCACTGGACAAGGTGTGCGAAACGCTTAGGGGTTCTGCTGACCCAGGTTTGTGCGAGAACACCAACGATATCGCGCC
GATCGGTTGGGTTGTGCGACCTGCGCTGCGATCGTAGCGGAGTCTGCGGTAGCTTCCGCGTACCTGCGCGCATGCTGCTTC
ACCCGCGTGACCGAGCAACTGGCCAGCGCGGGGATCAGGAGGCAAGCAAGGAGTTCTTCCAGGGGGCAGCGATCAGTTCTTG
CGCTCCTTGATGATGATCAGGGACTGCGGGCGGCTGGCGTTTACCAGGAAGGGCGGTTCCGCGAGCGTCATACAGGGGGAACAGC
TCGGATACCGCGCTCCTGAAGGAGCGGTGGAAGGTGAGTGAGCCTCGATCGGATAGTTGAGATCCTGCGGCTCCAGCGGACG
GTCCAGCGTGGCGCTTTGCCCAAGCTTCCAAGGTGTCGCTAGCGCAGGAGTTGATCCCGTCTCGGCTGACCGGCTGTCGCGGCG
CGGCTCGGTTCCGGCGCGGTGAGCACTTGAACGGTGGGCGGAAGCCACGGTGGTGGAGTGGGCGGCGTGGACTTACGGCG
GAGGCGCGAGGCTTGTGCTGGTGCAGAGCGCGGGAGGAGAAGAGGGGTGGCCATCCTTGGCCGCTGCGGCTGGGGGAGGACA
TGTGTGGATTGAGGCTTCTTGGCCGAGCGTGGCTGGTTGTGATCGGCGCGGTTGTCGGCACCGAGGCGACCAAGTGGGCGAC
TCCAGCGTGGTACTTTCTCCGTGGAGAGGTTGCGCTACCGCGCGGAGTTGGGGCAGCGGAGTTGGGATCAGTTGCTGG
ACCGGTTTCCGCTCGGCGCGGGGGAAGTTGATAGCCCGGCGCAGCAGCAAGCAGACCTGCCGCGCGACCTCGTCGACCTTA
ACCTGCCAGGCTGGGCGGAGAGGACCTGGAGGGTGTGCGCAGGGTTCATCGGCGCGAGCTTGTACTGAGCTGCGCGGAGCGGC
CGGGTGTAGAGGATGTTACATGACCGGCGTCTGCCGGGACAGCGAGTAACCCGAGCGGCTACACGTAAGTGTGCGGCTG
TTGACCGTTCGGGTTGATGCTCGACGGGATGGTTACGTCAGATGATCTGGGCCATCAGGTGCGGTTGACCGGATCAGGCTGGGTG
CTGACCAAGGTTGATGCGCCATAGCGCACTACGGGCTCTTCTCCGCTGACACCGTTTGGATAAGGTCGGGAGAAAGCAGG
TCTGGATTGCGGCTGCTGTGAAAGGCTGGCTGGCGGCTCAGCCTTGTGGGTGCTGAGCTGGCGAGCAGCAGGACAGCGTG
AGAGTGGTCAACTGCCTGGTCAATTGAGGCCCCCTTGTATGATGGGGCACAAGTGGCAGGCCATCCGCCACGATCAGCAGGAA
ATACTCGGCAACAGGGGAGGGGATTTCTCGATGACTAGGCGAGGTTGAGCGAGGCGTAAAGGAAACGGATTAGGAGCGA
TGAAGGTTTTTTCATGTTACGCCTAAAGCGAGCCAGGTGGGAATCGTTGTTAGAAATGGATAAAATAGAAGTTCTAATCGTTTTA
GAATGCTTTGTTGAACTCAGGTGAGCAGTACGGGGTGTAAATTGAGCTTTAAATACTATTGGGCTAAATTTTTCTGGGAGCT
TTCTTCTTTGTTTTAGTCGCTTGGAAAGGCTCCGTAATTTCAAGCCTGGCATCAGTTAACCCCTTGGTAGTGGCTGGATTGCT
ACTACTGCTGTTTTCTTCTCGGTAAGGCTTGTGAAAGCTTTCGCTTAAAGATATACGGAAGGAGTCTGGGTACAGGTTT
TTCTCCGAAACCCCTGCAAAACAGGATTGTATGCAGTCTTTTATTTGGCTTGTATTTGTTTTCAATTCCTTGGGGATGATT
TTTTTATTTCTATAAAACGGAAGGCTCGTAGGATTTAACGAGGCTTGTGTTGTTTTAAATGGACATTACAAGGTTGTTTGC
TTTTTCTAGAAGGCTTCTGTCATCATCGCCCGGTAAGTCCATTACAGTGGCAACCCAGGATGGCTATAGGAGTGGCGCT
TGCCGTGGCAAAATGCAATACCCCAAGCCAACTGGCGGCGCAGGACAGTGTGTTTCAACTTTTACAAGAAATGGTTTCCA
GTCCCGGTTTCCGTAGATATCTGAACTCTTGGTACAGGCTGGCGGCGTCAATAGCTTTGCCCTACAACCTCAAAGCCTTTGCT
AAATTTCTCAAGGCTCTTCGCCATCATCTGCTGTCTAGAGAATCAAAGGCTTTGGCAATGGCTGCTTCTTAAAGGCTAAG
TTTTTTATTTAAACGATCCTTGTGTTTTTCAAACGAATTGATTGCTTCCGCCGAACCTCTGATATTTTCCCCCTGGCGCCTTC

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND
PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September 12, 2003 Serial No.: Not Yet Assigned

Page 13 of 118

Customer No.: 21559

GGCCAGTTGGTGCGCCATCTCTGATGTTCTGTCGCCAAATTTCTCAGTTACTTCCCTATAAAAAGTCGGCAGTAAATTTAACAGC
ATCTTGCAACGCCTTGTCTCGGCCGCCTTGCCTTCTGCTTCTGTCTCGCTCGTTCGTTTCGGCGTTACAGTTATCTAAGCG
GGTCTGTAGACTGGCTATCTCATCCACAGTAGCTACCACTTGTAGATGGCTTTCTGCTTTCTAGAGGGGTGGGGCACTTCT
GATAAGAGTATTGTTTGGTCTCGACGTGCTTTTATCAGCAATTGCTTGCGTAAACAGCTCTTGTGGCAGTTCAAGCACTTG
CTCTGGCGTTGGACTGACCAATGGAGGCAAAGCACTTCCCGCTTGATCAAGCTCGGATTGACTTTTCATCTCGAGATTCCGCCT
TCTGGTTTTGTAGTCCCGCTCCAGTTGAGAGAGAGGTCCTCTTTTCTTTAGCGCTTTTGTGCTTGTGAGCTCCTGTCTTAGGAT
TTTTACCGCTTTCAGAGAGTGAATCTGTTTGTAGTAGGCAACCCGCCCTCCAGTTTCATCTCTAACAAGCTGCTGAAAGGGCA
CTTCTTTCTTTCTTCCCAAGAGTTCTCTCCAGATCTTGACCCAGGCTCATAGATGGAGAGGTTGACTGCTTTTCGTAGTC
AATGAATGGGCCTTTGGTAAAGGCTTTTTTGTATCCTTCAAGATCAGATTCCATCTTGTGCAAGACTGCCGACATGACATCGAC
TCGCTCTTTCGGCTTGTGGCCTCCCTTTTCTTGAAGTATCTGCTTTAGCAAGCCATCGACTTGAATCAGATATTTCTTGTGTAT
CTCCTTCTTTCTTTTGTCTTGTTCCTTCAACATGGATCGGATCGGATAATAAATGTTTTTCTGTAGATATATGTCTCTTTGAA
CAAGTCGCCAAAATATTTCCCGCAAAGCCGCGCTATCGATAGGGGAACTCTGATTCGCTATTGATTCGCTTGGACCTTGTG
GACAATAACCATGGAAACAGGTAATCTCGTTGTCATTGGACATTTAGACTTTCTCCATTTGGTGAGTGTGGTACAGAAAACAGTGGC
ATTCAGCTCTAGAAATGAAACCCATGGCTTCAATGTGTGGGCAATGCACTAAGAAGGCACGTCATGCGCTGTGATATAGTGG
FAATAAGCAGCGTTGTCTAAACTCGATATTCAGAAATGGGATGGCTGAGCCACGGGATATTCCTAAGAGCTCATGTATCTCT
ATAACAGAGTTTCATGCGCTCAAGATTCACGCTGGGATGTGATGCTCTCTGCAAAATTCAGTACTTGTATCGGCAACACCATCT
AAGGGTGAAGATAGGGATAGCAAGCGAAGAGACAGTTTTTGTGCGTTTCCATGTTAAACGGGAGATACAGCTCGAAGACCTAAGCC
ATTAACACGTGCTGACCATCCAGTTGGGGCCCATTAGGATCTAGCACAGCACCCATGGTGCACCTGGGGCTATTTCTCGGATGTA
CACCGTTTGTGATCAGCTCCGGAGAGACAGCTCTGGAATGCGGCTGCTGTCGAAGGCTGGCCTGGCCGGCTCTGGCCTTGTGGG
TCGTGAGCTTGGCGAGCAGGCAAGCAGCGTGAGAGTGGTCAACTGCTGGTTCATTGAGGCCCTTTGATGATAGGCACAG
TGGCAGGCCATTCCGCATGCTTACAGAGAAATACTCGCGGAGGATCAGGGCTTTTTTATCCAGTTAGGCTTGGCGCTAAAG
CGCGCGGTTTCCGTCGACCTGCTCGGGCTCGATGAGGTCCTGGGCAACCTGGCGTTTCGCGTTGAGCAGTTGCCACAGTTGCTC
GTCGATCTGGCCTGACCACTGAGGATTTTCAACACGACCATCGGAGCTGGCGGTTTCGGTACCGCGCGGTTCTCGGCTTGTTC
TCTGTACCGGGATCAGGGCAGGCCGAGGAAAAACAGTAGTTGCGCGCAGTGAGTTGTTGCCGCTCCCTGCGGCCCGCT
AGTGCAATGAACACTCGGCAGTCGGGATCTCTGGAAGCGATCTATCGCCTTCTGCGGCTTGTGGAGCAGTATTGCCAC
CAGCGTGACGACCGCTGTCCGGCTGCTCGCAGAGTTCCTTCAGCGCAGCCACGGTCGGCTTGAATCTCGCAGAACAGGATCAC
TTGTGCTCTGCGCTCAGCTCGCTCAACAGGTCATCGCCACGCGAACTTTACCGTTTTCAGAGTAAACGCCGACGCGCGCGAG
CTCGCGCAAGACCGGTCGGTCTCGAGGCGAGCAGCTCGTATTGTCGGCTTCTCGGTGAGAGAGGCCACCTTCAGCAACTG
CCGCTGCTTGGCCTTAGGCTGGGAGCAGCATCTTTGCGCTCGCGACATCCAGTACCCAGCTCGCGCCGAGACTTGTGGCG
GAACTCCGGGTTGCCGCGCAACCGTTCGAGAACTCTTTACGCGGAGTTGGCCGATGGGGTGGCTGAGAGGCGCAGCAGGCT
GTGAGCTCTGTCTCGCGGTTGAGCACCGGCTGCGCGTAAAGAGTGGGTTCCGCACTTGGCGCGCAATGTGCAAAACCGTG
CCGCTGTCATTTGCGGCTCGGTTCTTCTATCCGCTCGGCTCGCTCGATGACCATCACGCGAGCGGAGCGGAGCGGAGCA
GGGGCTCAACTGCTCTGATTTGACGAGGATCACTGCGCTCTGGGGTGTCTGCTGGATGGCCACGGTGGCGGAGGATAGAC
TCTCTGGATCTCCCGTCCCAATTGATCAGCAGGATAGCCAGGCTGATGACAGGATTGGTCTGCCCGCGCGGATCGAAGC
GGCGATGACCGCCTGGCGGGTCTTGGCCCAATCCATGCTCTCGGCCAATAAGGCGCTGGTTCTCTGACGAGATGAGCGATGCC
GTACGGCTGGTGGGCGAGTAGAGATAGCCCTGAAGCGCGCCTCGATACAGCCGAGCTGTACTCGGTGCGCTCGATCTCCG
CACCGCGCGAGGTTAGATGTACAGAGCTGACTCGTCTCCAGGATGGCGCTGCGGTTCTGCTGTGGACGCCGATGTCTCAT
GCTGGGCGAGCTCGGTGACGCGCGAAGCTGCGCTCGCTGAGCAGCTCCTGGACAGTATCCAGAATCTCGAATTGTTCTCGG
GAGGCGCAAGCTCCAGAATCAGGTTGCTTCCGAGCAGCTCGGAGTGGTGTGATGCGCCAGGAACTCGATGGGCGCAGGAATAC
CCCGCGCATCCGCGAGCCACCGGCAACACACAGGATGGAATCGCTGACAGTAGGACCGCGCCTCGGCGAGGGCGCGAT
GCGCAGCAGCATCCCCAGGTGAAGGCTGTGCGTCAAGCGAGGCTGCGCGGATTGCAACCAAGGAGCTGAATCCCCGCAAGGA
CTGGGAGTGAAGGCTCCAGCTAGCTCGGCGAGACGGTTGTAAGAGTGTCCAGTTACGGCAGCAGCTTGCTTTGGGGACCG
CCAGTAGCCGTTGCAACCGATGCTCCCGGCGCTGATGAAGAAGCGGCAAGCTCCAGCAGCGCGGATGCTGCTGCGATGTA
GGGATCTTCAAGCCGAGTCTGATCGGCTGGAAGTACCCGCTGCGAGAGCTCGAAGGTGTCGGGAGGTACGATTAACCA
GCGTGGACGGTTTCATCGGCTTCCAACTGAGTGAGCAGCACTTCCGCGTGGCTTGGCGAGGCGCCCTTACAGCAGCAATTG
GCGAAGCGCGTCATGTGACGGTTCATGTGGGCCAGTTAGGTTTCTGCTATGGCTCCAGAGCGAATTGCTTCCATGAACCA
CGGTATGCTTGGCGCGCAGGCGAGTCTGGTCGCGCACGGGGCGATTGCAATTCGCGCAGTTGCACTGGAGTGAAGGTTGTTGGGA
CTCCAGGCGCTACCTATAGGCTCGCATGTCTCACGCTGAGCTCCAATTCGGCAGAGGCGCAGCGACAGCCAGCTACGTA
GAGAGCAGTTCGCGAGTCTCTCGTGGTGGCGCTGGTGTGATGATCATGATGCAATGTGCGCAGAGCGATATGTTGCGCATATAG
TCGGCAGGACAGCGGAGCTCGACGATCTGCACTCCATTAGGGCAATCGATAAGACCGACAGACAGCATTTCCCTCCAGGTGGT
TACTCTGTCGGTGTGCAACCGATCGGCGAGGCTGCTCTGCACTCGGCGACGCGATCCCTCTGGAACCTGGTGGTAGTCTGCTCCAC
AGGTTGCCAGCTTGTCTGGTAGGTAACAATGACCCAGATACGCTTTACGGGCGCTGCGCGCTCCGGAACACCTGCTGATTTCC
TGGTTCAGGGCGGTTGATGAATCTTGTATGTTCTCTACGCGCTGAGTGATTGCAAGCAGTCCGCGATCTGCGAATCCGACGG
CAGCCGTTGAAGAGACGGTTGTAGTCCGGCTTGGCGTTGTGAATTTGCCAAGGGATCGCATTGTAGGCGGCATAGAATGCTGTC
CATTGAGCGCGAGTATGTCGGCGCGGTTACTTAGCATGCGCGCAGTAGGAGAGCATGCCATCCGGAATGTGCGGCTTTCCGCT
TTGCGATGGAAGAGAGTCTCTCCGCAACGCGCGCTGACGTGACGTGACTAAGAAATCGAATCGAGGCGCGCGGCTGAAATAGC
CAGGCAAAAAAGTTGATGACCGAAATCTTCTGTCGCGCACTGCCAAGAAATACGGCCATATCGAGGTACAAGTCGATGCG
ATGGCGCGCGCTGCGGTCGTTCACTGTCAAGCTGGGGAAGAGGGCAAGGAAGGTAATGGCGGATGCTCTCGCGGTTGTGCTG
GTCGTGCTGTGAGGCTCAGACTCCCTCTCGCAATCCACAGTAGCGGCGCTCAAGACCGGCTGACCTTTCAGATCTGCGACCGG
GCCACCTGGTGGCTTGAAGCCAGCCAGTTGTAGAGCTGGGCGAGAGGCAACGACCGGATGGGCAACAGCAGGCTTGTG
TCCAGGACTGCTGTGTAAGGTTTCGCAAGACGCTCCGATCTGCTTCAAGTAGAACTGGTCTTGGCGCGTAGTAGCGTTGTG
CTGTGGCTCGGATACACGAGGCGATGTTGAAGTCCGAGATCGTCCCTGGCCATCTGTGTGGCTAGGTCTGCGCTCGCTTTGTG
TATGAGCTGAGTACAGCAACTGCTGCTGCAATTTTGAAGGCGCTGTAGCAGTCTCAAGGCCCTGTGTCGAGACGCGCTGCTT
GCGGGGATACGATGGGCAAGGCTTGCACTCCACGCAACCAATTAGTTGACCGTCAATACCTGGAGTAGCGGCTGACGCAATG
ATGGTGAACGCACACCGTTGTGACGAGCAGCTGTGATGACCAAGCACTGGTCTGATGTGCTGCGCAGGTAGTACCGGAT
TGGGCTGAGGCTGAGTGTGCGGCGCGGTTACTTAGCATGCGCGCAGTAGGAGAGCATGCCATCCGGAATGTGCGGCTTTCCGCT
CAGCCGTTGAAGAGACGGTTGTAGTCCGGCTTGGCGTTGTGAATTTGCCAAGGGATCGCATTGTAGGCGGCATAGAATGCTGTC
CATTGAGCGCGAGTATGTCGGCGCGGTTACTTAGCATGCGCGCAGTAGGAGAGCATGCCATCCGGAATGTGCGGCTTTCCGCT
TTGCGATGGAAGAGAGTCTCTCCGCAACGCGCGCTGACGTGACGTGACTAAGAAATCGAATCGAGGCGCGCGGCTGAAATAGC
CAGGCAAAAAAGTTGATGACCGAAATCTTCTGTCGCGCACTGCCAAGAAATACGGCCATATCGAGGTACAAGTCGATGCG
ATGGCGCGCGCTGCGGTCGTTCACTGTCAAGCTGGGGAAGAGGGCAAGGAAGGTAATGGCGGATGCTCTCGCGGTTGTGCTG
GTCGTGCTGTGAGGCTCAGACTCCCTCTCGCAATCCACAGTAGCGGCGCTCAAGACCGGCTGACCTTTCAGATCTGCGACCGG
TCCACCTGGTGGCTTGAAGCCAGCCAGTTGTAGAGCTGGGCGAGAGGCAACGACCGGATGGGCAACAGCAGGCTTGTG
TCCAGGACTGCTGTGTAAGGTTTCGCAAGACGCTCCGATCTGCTTCAAGTAGAACTGGTCTTGGCGCGTAGTAGCGTTGTG
CTGTGGCTCGGATACACGAGGCGATGTTGAAGTCCGAGATCGTCCCTGGCCATCTGTGTGGCTAGGTCTGCGCTCGCTTTGTG
TATGAGCTGAGTACAGCAACTGCTGCTGCAATTTTGAAGGCGCTGTAGCAGTCTCAAGGCCCTGTGTCGAGACGCGCTGCTT
GCGGGATACGATGGGCAAGGCTTGCACTCCACGCAACCAATTAGTTGACCGTCAATACCTGGAGTAGCGGCTGACGCAATG
ATGGTGAACGCACACCGTTGTGACGAGCAGCTGTGATGACCAAGCACTGGTCTGATGTGCTGCGCAGGTAGTACCGGAT
TGGGCTGAGGCTGAGTGTGCGGCGCGGTTACTTAGCATGCGCGCAGTAGGAGAGCATGCCATCCGGAATGTGCGGCTTTCCGCT
CAGCCGTTGAAGAGACGGTTGTAGTCCGGCTTGGCGTTGTGAATTTGCCAAGGGATCGCATTGTAGGCGGCATAGAATGCTGTC
CATTGAGCGCGAGTATGTCGGCGCGGTTACTTAGCATGCGCGCAGTAGGAGAGCATGCCATCCGGAATGTGCGGCTTTCCGCT
TTGCGATGGAAGAGAGTCTCTCCGCAACGCGCGCTGACGTGACGTGACTAAGAAATCGAATCGAGGCGCGCGGCTGAAATAGC
CAGGCAAAAAAGTTGATGACCGAAATCTTCTGTCGCGCACTGCCAAGAAATACGGCCATATCGAGGTACAAGTCGATGCG
ATGGCGCGCGCTGCGGTCGTTCACTGTCAAGCTGGGGAAGAGGGCAAGGAAGGTAATGGCGGATGCTCTCGCGGTTGTGCTG
GTCGTGCTGTGAGGCTCAGACTCCCTCTCGCAATCCACAGTAGCGGCGCTCAAGACCGGCTGACCTTTCAGATCTGCGACCGG
TCCACCTGGTGGCTTGAAGCCAGCCAGTTGTAGAGCTGGGCGAGAGGCAACGACCGGATGGGCAACAGCAGGCTTGTG
TCCAGGACTGCTGTGTAAGGTTTCGCAAGACGCTCCGATCTGCTTCAAGTAGAACTGGTCTTGGCGCGTAGTAGCGTTGTG

15/118
FIG. 2M

AACTTGGTTTCGAGCTGTCCCTGCGAGGGAGAAATGTGTTGCCAGGTGTTGGCGACGGCGAGCTGCCGCTCGCTGTAGCGGGC
AGGGATGGATCCTGTAGCGGCCCGCTAGGGGGGGCGTTGTCGAGAGGTGCCACGCGCATGCTCCTGTGTAAGCAAGCAGCACTT
TGGCCCGTTCGGTTTACGCTGAATCCATGTGAAATGGTCACTCTCAAGCGTGTGGATTTCAAGGTTCTACAAGCTAGCTGACTGA
AGGTCTTATTGGGATGGATCAGTAAATTTCTCTGTAGTCAGCCTCAAAGTCTCCTAGTGCAGCAAACTCGTCAAGGATAGCTTT
ACTTAAAGAATAAGATGTATTGTCTCGCTGGGTGTGCCAGTTGGCACCACCATGCTTGCCTCCGTGAAATAGATTGTTTCG
CACGGTCTTAAGTAAGCGAGCGACTCTAGCTAAGTCGCTTTTGTCTCATCTAATTTAACCGGACGCCAACCCAGCTCTCTACC
GGGCAGGACTATTTGTGTCTCTGACTCTGTCTCGATTAGTGTGTGGCTGATTGGGAAAGAGAGTATTTGTGAGAATGGTTTTG
TACAAAATTTTCCCATCCCGGCTCTGCCCTAGCTCCAGGTTTGTAAATTTTAAAGTAGCCATTCTCTTTGAGGCTGAACCTCAA
TCGCGAGAACCAATAGAAAAAATCAAAAGCTCGCTCTGTGATTTCGGCAGATATATCTCTTTTCGCATAAACCCCTCCTTGATA
GCTAGCAGCGAATTGGGCTGTACCATTTTGCTGCACTGCGTGAATACTATAGTGGCTATAGCCCGCGCGGGTTGTGCAAGT
CAGGCCGGCCGCTCTTGGTGGGGTAGCTAACCTTGACCTGTGTTGTGATAGCCGCTGCAACCCAGAAAGTGAAGCCGCTT
TGCCGGCTTTCTTGGGTGAACAGGGGCGAGCCGCACTTGCCGCAAGGAAGTGCCTGGTGGCGCCGAGGAGATACCTGGCGG
TGGAGCTTGGCGGTGTTCTGTCTGTGCGTGGGGAGCGCGGAGTGTCTGCAGCTCTGTCTGTATAGTGGCTATAGCCCGCGCGGGTGTGCAAGT
GGACCACGTTCACTAGGTGTCTCGCTGTGCGGATCCGGTCAAGGCGACCTCGAGGTTGGGGTGAAGTGAAGGCTGAGGA
AGCTGAAGTTGTGCTCAGGAGCGGATGGTGGCTCTCCAGCTCGCGGGGATCAGGCTTCGCCCCCTTCACTGACCTGGC
CCTTGTGATGATGTTCTTACGATCGAGGCGTAGGAGGATGGCCGACCAATCCCGCGCGCTTCATCTCGCCACCAGGCTGG
CGTCGAGAACCAATGAGGAGGGTGGTTTTCTTCCAGGACTCGCGCTGTAGACTTGAATCTGCGGTGGCTCCAGCG
CCGGATCGGGATGGGGTTTGAAGGCGTTTCTTCTTCTCGCTCTCGCGCTCATCACCTGACAGGTTTTCTCCAGCCAGGCA
CGTTCAAAGCTTCCCTTTGGCGCCGAAGCGCAGCGGCTTTTGTGGGGCCGACGCCCAGGAGGGTGTGTTCTCACTGCGT
ACACGGCAGCTTCGATCTGACTGGCAAGCGCGCGGACTCGAATAGCTGGTACAGCGCCTGCTCATCAGCAGTTTCACTGGCGG
CAGCGGCCATCCAGTCGGTAGGGTGTGCGGGGTGGCTTCTGGGCGTCTTGGTCCGCTTTGAACATCCGCTGTTGCTCAA
CACACTTCAGCCCAAGGCTTTGGCGCAGCAGCGGATCTCCGATCGAGTCTTCGAGATATTTGGGGTTGTCCGTCGGTGGT
AGGTGATGAGCCCTGTTTATACAGCCGCTGGGCGACCTTCTGCTTGTGCGGTGACCACTTCAAGCGCATCCCGCGCGCCA
TCTGGAGCGAGGAGGAGATGAACGGTGGCGGAGGCGCATCGGTTTTCTTGGCTATCAATGCAGGTCTCGACGATGACATTACGTA
TAGCCCCCAGCTTCTGCGAGTTGACGATCCTGGACATACGGGAATCCTCGCTGGCAAAACAGGGCACTGGATCCCATTCG
CCGTCAGGTAACGGCGCTCGCTGGGCGAAACGAAGGTCACTTCCACCCGAAGTGGTTGATTGCTTGAAGGCGCGGATCTCC
GCTCTCGCAGGACCAACAGGTACACCGCGACGGAATGCAAGCGCCCGCGGTGGTGGCTTACCCATCACGCGCGCAACTCTG
AGCTGAGTTTCGGCGGTGATGCACTGATTTGTGATTTCTTGAAGCAACGCGCTTGTAGTTCTGATCCCGGCGAGCTTGGAGGA
TGTGCGAGCAATGCTTTTCCGCTTCGCGATCCGGGTCCGATCGCAGTATAGATTCCACGCGCTTTCGCGCGAGCTCTTCCG
GTGCGACGGTTTTTTCTTGGCCGAGAGGATCTGGTAGTGGGTTTGAATCCTGGCCACGCGACGGTGAGCATCTGCGGAT
CCTGCCGTGAAACGGTAGGTCTCTGATGTGCGCGGCTGTGCGCTCGACCTGCCAGGTCACTGCGGACGAATCGAGGGGAGAA
GGGACCTCAGCTTTTTTCACTTGGCCGGGCTTCAATGATACGATGTCATCGGCGGAATCTCTCGAACAGTCGGTGAAC
GTGCGGATCAGAGAGCGGCGCAACAATGCGAAATAGACGGTGGCTTGTGAGTTTTCGGGGCAACTTTCGCGCATCTCGGAA
AAAATATCCGTTGCCGATCAAATCGGACACAATTAACGATTCTCAAATGCGGTAATCTTGTCCGTGTGGTGTGCTGCAACCAT
CTACGTAATCGTGGCAACTGGGTGCTCGCTCGCATAGATTGAGCAAGGAAATGATCGTTCTTGAAGGACATCAATCGACCC
CAATCATCCGATTTGGTCTAGGAAGGAAAAATGGTATGGATCAAGGCTTTGACATGCAATGCCAACGCCAATCGTCAACCCCA
AGGAGTGGCGACTGTGCACTGTTAGTGGTAGAACTTTCCGCAATTAACATTGATCGCAGGCGACCATTTGGTTGAGCTAGCG
GCAGCGCTGGTGGGTGATGCGGATGAGCCCGCAGCGAGGACGAGGTGGCGGCTCTGTTGGTCAAGGCTGGTGGTGTACTA
CGTGTGTTGGTGGGATAGGAAGCGCATCAGGTACACCAAGCCCCCTCTAATCGTGGGGCCAGGCGGTAGGTGAGCCAGC
TTACTTTGCCCAATGCGAAATGGAACGGTGGCTGTTGAGTCTTCGGGAGGAGCTGAAGGTTAAAGTCACTAGAAGTCGAGC
TGCTTGTGCTGCTCTGGGCAAAACCTTTTCAACAGGACGCGGATCGCAGGATGACTGTAGGGGCACTGGAACACCGGTAGC
CCTTTCCGCGCGGCTCAGCACAATAGCCCGTACAGTCCGGTGGGATCGTTGAAGTGTGCTGTCTGTCAGCTTGTGCGTATCG
GTCGCTCTCTTTGACACACTGCACCTCGTACTGGTTCGAGATTTCGATTTCGGAAGCCCTTGGCACCAGAGACTTCAAGAACGG
CGGCCAAAATAAGGATGCTGGTCTATCAGCTTTTCCGCGCAACTGTTTATTGCGGTTGCTCCAGTCGCTGGAATGCTTGTG
GAGGCTGCGGGCGCGAAGAGAGGTTCTTGGCACCTGCAACTTACAGAGTGGGCGGCAAAAATAAGACCGGCTTTCACCGTC
GCGTCGGCAATGCCTTGGGCCATCCTGATGAAGGATGCGCGATCTTCTTCTCCATAAGCGCCATGCACTCTAGAACTGATATC
TCGTGCGCGTGTAGGAGCTGCCTGGTCAACGGCTCATCGCATCCCGAAATGAGCAATGTGAGCCTGCTTGGTCACTTTGCTG
TTCATGGCGGAGTTTCTGCGGGGGCTTTGTGCAAGCTGCGGGTGGCAAAAGTGGTACTTGGCAACCCGAGCTTACGAATG
GGCGGCCAAAATAAGCGCGCTGCTACATGGGATGTGCTCGTGAAGGCGCTGTGCTAGTCTGCTGCGATTGCGCATCTTG
CTGAGCAGGTTGTGCGGGACGTTCCGAGACTGTTGATGCTGGGTGGCAACAGGGTGACCTCGGCCAGGCGATGCGGAAGGAT
GGAAATGCGCGAGGCTTCGACCTTCAGCGCTGGAATCTTTCGCTGACTCTTGTCCGGCCAGCGGTCCATAATGGCCCTGCC
TTCGACCTTGACGCGCATACCTTTCGTGAACAGTTTCGGCAAGCGCTGGGCATCCTGATGCCACCATTCGACGCTGCACAGAA
GCCCGCGCATCTTGTAGCCACTTGGCCATCGGGAATCGAGTTGTGCAACATCACGTTCACTGCGCAACTGCCCGGGGG
ATTGTTGCCGTTGGGGAAGCTGCGGTGCTCCGGCGCGGAGCCAATGTTGGCTTCCAGAAGACGGGGTAGCCATGATGAATCT
CCAGTCAGTGGTGGTGGTGTCTGCTGCACTGCGGCCATGAATCGGCTCGGCTCGGCTGGCCATCTTTTCCGCGCACTCGGCGG
CCTGCTTGGCGGATGGAGTGGATCAGGCTGATCTGATGTTGAGGGTGTGCGCTGCACTCCATCTCGTAGAGGCTTTGTCAGCA
GGTATTCCGGGGTCTTGTCTGCGCGAGCATTTCCGTCCACACCCGAGGCTTCCGCGGGATGCAATGCTGCTGCTGCGCAAGG
GAAAGTTGTGCGGGCGGATGTGCGCTGCTCGATGAGTGGGTGGGAGTAGAGTGAAGGTTGGGCGACGAACCTGGGCCAGCA
GTCCCTGGGCAAGTTCAAGCAATCCCTGCTCCATGGCTTCGCACTGTTCCGCCAACTGGGCGAGCTCCCTATTACCTTTAAAG
GCTTTAAAGGCTTTTAGTGAGGCACTGTGTTGACGCGCTGGAAGGCACTTGTGTCAGTGGGCGGAAGGCACTTTGTTCCA
GAGCGCTTGTGCTGCTTATCGAGCTGCGCATTCGATCAGTCTGAGTTCGAGTTCGACACGCTGATCATCGGCAACCA
GTCAGAGCGGCCCCGGGTGATCGGCGGAGCGAAGTTCGAGCGTTCGAGTGCCTCCAGGATGTCTGCGGGATCTCGCCGAACCT
CTCGTACATCTTCCGCGCGGCTTCGGCGCGAGCATTTGTTGCGGCGAAGTGTGCGGAGTGGCGCCGAGAACTGGTAGCTCTG
GGCAGACCAACAGGCTTCGAGCACCGACGCACTTCTGTCGAGCCAGACCTCATGTGCGGGCAGCCGACCGAGCCGAGCGT

16/118
FIG. 2N

CTGGGCCAGCAGGATTGACGCGGCGAGTTTCGTCTAGTTGGTCAGGAGATAGACTGCCTTGAAGCCGAGAGGGTTGGAGATGAA
 CAACGGGACCTTGACCGGTGTAACGGACAGGTTCTCGGAGATATCGAGGGTCGCGGGCAGCTTGGCCATGACGTATCCAGTCG
 CTGGTCGATGGCCTCGAGTGGGCTTGGCAGCTCTCGACTTCTCTTCGATCTGAACAGCCACAGTCGGAGTACGGATCGTC
 CTGTGCTGCCCGCGGTGCATGCGATTACGATATGCGCAGAAAGCCGAGAGGCCGAGGATGCTGTGCTTGGCATCCGACTTCTG
 CCGGCCGGTCCAGATGCGGGCGGCTGATAGGTGTGAAGCGTCAGAGCGACATCGCTGCGGAGAGAGCCGATCTGAAGCCGATG
 GGTTTCAGCCATGGTGTGTGTTCTTTATAGGAAGGGGTCGCGCCATCGTGCTGATCGGGGTAGGAGCGCGCAGTAGGAA
 ATACATGGTTGCGCGCTGGGGAATTTCCCATGGCGACTCGGAGCAAGGCACCCAGTTGCGCGCGCGCGTCGAAGTCAGTTCCTA
 ATGGAGTGGTTCCAGTGGAAACCACTCGCGCTCGGCGCCGAGTCAGTTCCTGGTGGAGTGGTTCCAGTGGAAACACTCA
 TCTTGGCGTGGCGCCGACTTGGGTATCTCTACCGTTGGTAACTCTGGTTGGCAGTTCATGTAGTTGGCTTTGTTGCATAAAGCT
 CCCCACATCGCTTTTTCCGCCATTCAAAGCGATACATCTCAACCGTAATGGTCAACACGGTTACTCGGAATACCCGAAACAGATC
 AGCATTAAACAACAGCCTGAGTTCTCGGTCTTCCAGCATCTGTAAAAAGTGCAGAAAGAACGAGAACGACCAAAAAGAAAGGG
 GCCATGCTCAAGAACAATTGTTTGAAGTTCCTCGTGTGGTGTGATGCTCGCTTGGCGAAGTGTTCGGCCCGGAAACACCCATCT
 TGCCGCGTACCGAAGTGGATCCCTGGCAGAGTATCCAGTGGAAATCAACAGCTCGCGCGCGCGCGGAGGTGCGTCTCTG
 GTGGAGTGAATCCAGTGGAAATCAACAGCTCGCGCGCGCGCGGAGTCAAGACCTGGGAGAGTGTGTCGGGAGTGTACCAAA
 AAAACACGACGCGCCGAGGAGACTTCTCATGAATCATACCCCGCTCGGATGGCAGCTCGATCCGTGCTGGGCCGAATGAG
 GTCTTGGAGCGGCTGAGGACCTCGCTCCCTGTCCGGGTATCCCCCGGGGCTGGACCTGAGGCTGTGTAGGCGCTATGGGGCT
 TGATGGTCCGGAGCAGGAATGGGCGCTTCCGTTGCGGGATGGTTTCGCTCAGCCGGTCTTCCCGAGCCCAAGAAGCGTTGAA
 CTGCGCTGAGCTGCACGCTGGATGAGCGTCACTGAGGTGAGGTAAGCAAGCGATTGCTGACTGTTCCGCGGCAACCGGATCGCA
 CTGCTCGATGATGGGTAGCCGAAGATCTCGGACACCCGCTGAGCGCGCTTACTGCTGGATCCGTTGCTCGGCTCCAGTGC
 GTGCAGGAGATCTGCGAGTTCGGATGGGCTTCTCGCGTGTCTCGCGGTGTACTGGCTTTTACATACAGAATCTTGTGTTGTA
 TGTATACGTACTATGATGATTCGGCATCCGAACCAAGGCGTACTCGCGGATTCGTGTCGGGTTCGGTATCCGAATCAGCGG
 TAGATGCTGATGCTTCCCGCTGATTCTCACTGAGTTTCGGATCCGAATCCGAGTTCGGGGTGAAGTGAAGTAAATCCGACGAGTTCGGAT
 GCCGAATCTGCGCGTGGATCCGCTTCCAGCTGACTATCGATAGCCATTGCTGGTGGTTCAAGCGTCCCTCGAGCTTCTCCAG
 ATGGGTAGGAGGCGCGACCCAGCTCCGGATCATCCCTGAATCCCGCCAGATGATCTGCCCGATTTCCGCTATGGCTCGGTT
 GCGGTGACCGGTGGATTGGCTCAGCAACTGCACTGTAGTTCGGTTCAGCTCCAAGGCTTCCGCTGGAGAGACAGGCTCGTCTGT
 AAGGATCTAAACGTTGCCCTGGACCTGTCCGTTGAGTTCGTTGCGCAGGCGCTCGGCCGAGGCTCGGCCGAGGCTGAGCCGAAG
 GACCGTGAGTGCCTTGGCGATGGTTTCCGCGAGGCGATCTTCCCGGCTGCATACCGAGATAGGGGCGCAGTTGCTCGTATGT
 CGGGAAACGCGGTGATCCGCTCGTCTGATGAGCAAGCGGAAACCTTGCAGCAGTTCCGCTCCAGCGGAGTAAGACGCTTATC
 TCGAGGAGGCGCGCTGGAAACGTTTTCATGAGGATTCGCGCTGAAATGATGCCGGAATATTGTTGAGCCAGCGGTTCCGCGCT
 TCCGCGCAGCAAGCGCTTCTTATACAAATGGGCTTCGAGATGCCCCGAGGCGAGTCCAGCAGTGTGCGAATCGGAAAGCGCGA
 CAGCCTCATGGTGTGTGTCGCTTCCGACTGAACAGGTTTCAGTAGTTGTTGACGTTCTGTCTCAGCGTTCACTGCGTTGTT
 GCGGATGGGGCAGAGTGGGGCAGGTGTGAGCTGACGCTAGCGAGGTAAGAGTGGATTGGGACTTTTGCAGGCGCGCTGGCTC
 TGGCTCGATTGTCAGACGAGATAGAGCCGCTTCTCAACAGCTCTGAGTAGAGACCAACAATGGCGAGGCTGATGAATCT
 GGACTGCTTGTGTTCCGCTTCCGCGCTTCTCTGAGGATTCGCGATGTCGAGTGTCCAGCACTCATGTGTCTTCAAGCGG
 ACCTGATTTTCTGCATCAGGTGCTCCAGCGGAACATATCTGGCGCTCGATCTGTTTCATCGAGCTGCCGCCAGCGACCGCGG
 CGCGGGTGGATCTCCAGCATCAAGCGCTTGGTGGCGGTGCTCTCCGCGCAGACAGCCGAAAGCTCTGGATCATCGTTGTGGT
 GCTCCAGCTTCAGTGCCTTCGATCTGCAAGTCTTCTGCGCGCTGCGCTCGGCTGTGGCCAGGATTTTCTCATGACGTCG
 ATGTTCACTCTGACATCTACCCAGGAAACCGGGGTATTGCTGAGGATGCTCATGACCGCAGGCTGTTGGAATCGGAAAGCAATC
 TCCGCGCGCAGCCCCATCTCGGCACAACGCTGAGCTGGCCGTTGCGCAGGTGGTTGAGCACCTGCGTCAGGACCGCTTCAATTG
 ATCGTCACTTCGACATGTGATCCCTCCCCAAATCGTGCAGAAATCACTGGTCACTAGTGTGCTGCTCGAGGCGCGCGCGCGC
 AGCTCAAGGTCAACAGGCGTCGGGCAAGACGAATGATCCGGAACAGTTTCAACAGGCTTCGTCATGAGACGGGCTGCCTGG
 CAGGATCGGTGATCGGGTGCAGGCCACCGACTTGCAGTGGGATTGCTGCGATGTCAGAGCTGTTTCAATCGAAGACAGCTAGATGTTCC
 CGGAGGTCTGTTGATCGTGAAGTTCAGAACGACTTGGACCTGCCCGCTGACCGCTTGAAGCAGCGTCATGAGGTGCAGCGCT
 GAATCAGTAATCTCATGGTCTTCAATTGGGCTCGGATACGTGTAACCCACACCGCCCTGAATCTGACGATCTCGCTGGGGCG
 CCGACATGCAGAGCGATCTCTTCAGCCAGATCAGCGATGTGCTGTGCGAGGATCTCGGGGTATCGATCGACCGCTCGATGTAC
 CAGAGATCGGTGATCGGGTGCAGGCCACCGACTTGCAGTGGGATTGCTGCGAGAGCTGTTTCAATCGAAGACAGCTAGATGTTCC
 CCCTCGAGGCGCGCCAGCGCTTGGCGAGTCTGCTGGATCTTAGTCGATACCGGGCTCAGATATGGCTGCCAAGACCGCGCA
 CGCTCCTCATCGGGCAATTGCTTTGTTGGACAGGTGGTGGCGGAGCGCGCGGGGGAGGGGATTACCGGGTTGCTCCTGTGTC
 CTGGACGTTTGGGATCTGCGCGGGGCTCTGTGTTTGTGCTTGTCCAGTAGACGACGAGCAGGAGGTTTCTGCCCCAATTGGT
 GGGAGTTGAGTTGGTGGCGAGGTGTGGGCGAGTGCAGCAGAGTGGCGCGGCTGCTCCTGCTGCCGTTGGTAACTCTCGAGC
 AGGATTTGGTCATAACGACAGGCCAGGGGCGCTTCAATTGGTGCATGAGTTGCTCCTGGAACGCTCGAAAATGAATTCGTCA
 GGGCTACTGTGCAAGATTGCCAGGCTGCTTGGAAACAGCAATTCGAAGTCAACCCCTTCAACAGCGTATAGACGTGCCAGCAG
 GAGGATGCCGACTTTCTGAGTGACAGGAGTTTCTCGATTGCGGTTTACCGAGCCCTGAATACAGCATCACAGGAATCGCCGGC
 GCCAATACCTCAATGCTCTAAACATCTTACTGATAGGATTTAGATACAGGCTAGCCGCTTTCAGGCTTTCAGGACGTCGCAAC
 TCGCGCTGGGAAATGCTTTACCGCGCTTCTTGTTCGTAAAGAAATTCGCTTCTGCACCCCAACCGCGCGCTCGATGAAC
 TTGAGGTGCGCCCTCAGATCGTTCTCGGCCAAATGACCGGTGAGCGCGATGATTTCGCGCGCTGCTTGTCCAGGGCTTGAAC
 AGGCAGTCGAAGCTGAAATAGCGCTCGTCTCCGGTCTCTTGTAGAGGTGCTTGAAGATTTCCAGGCGAGTGTTCGCGCGT
 CGAATGCGGATTTGCTTCTTCCAGGCGCTCGAGTGCAGTGTGAGGCGCTCGATCGAGGCTCGATGCTGCTTTCGCTTTCAGC
 TCATCGTATTTCCGGTTTCAGAGTGGTCCGGGGTTATCGTGCCAGGGGAGGACGTCGTGGAGTGTGAGCCTCATGGGCGTGTG
 CTGATCGGGTCCGACAACTGCTCCAGAACCTTGGGCGGTGCGGTAACAGGGGCAAGTAGCCGCTTTCAGATCTTTCGCGG
 TTGATCTTGGCCATACTCATGCTCCGCCACCTGAGGCGGTGTTCCGCGCGCGCAGCTCAGCCTGCAGCGCTGGTTCTTGGCGA
 AGAGTTCGCTGCAACATTGCTTGGTTCGCTCGCGGAGTCTTCTTCGAGCGCGTGTGCTGCTCCAGTCCGCTAGACGATCTCGCTGT
 TGGCCGAACGGTGGTTATCCGCGGCTTTTGTCTTAGCTGATCGCGCAAGCCCAAGGGCATAACGACAACGAATTTATCTTCGG
 TTTCTCGGGCATGTCATTCTCAATTTGGGCGCTTGGCGCGCTTGAAGTGGTTCGCGGAGTTGAGTGGGAGGCTGTGATCTG
 CAGCGTGTGCGCTCCTCGTCTACTCGATCCTGGAGCCAGCAGGTGCGCCTCGAAGCTGATCGACAGGCTTCCGCGCGGCC

17/118
FIG. 20

GGTGAAGCGGCGAACTGGTTGAGGGTGCCTTGTCCGCGGGATTTCGGGCGACAGGCCGTAGTCTTGTATCGGATGTAGTCT
 GTAGAACGCCCGCGTTGCTGGTCCATCAGTTCCGAAAGCGCGTCAGGGGTATCGGCTCGCCGATGCGCGCCTGCAGGT
 GCGTAGTCGACACGCTCTCGGTCTTCTCGCGGGCTGTCTCGGCCATGTCTCGCTTTCCACGAAATCGCTGAAGGCTTT
 CAGCAGGGTGCCTGCTCGCTCGCGCAATCCACCCCTTCTGGCAGCCGATGAAGTCGCGGAAATAGTCGGAGACCTTCTCC
 GCCCTTGGCCTTGATGAACGAGATGTACTGCTTCGATGCTTGTGTGCGCCATTCCGAAATGTGATCCGCGCGCCAGGTG
 CAACTGGCCGAGGTCCAGGTGGCGCGACGGGGTGACCTCCAGCGACTCGTTTACCGCCACGCCCTTCGCTGTGGTGCAGCAGGC
 GATCGCCAGGTAGTCGGTCATGCCCTGCTGGTAGTGGGCGAACAGGACGTGGCCGCGCGTGAAGAGATTGGACTCCTCCATCAG
 CTTTGGCAGGTGCTCGACCGCTTCGCGGCTGAAGCCGACGAAGTCGCGGTGCGCTCCAGGTACTCGCCGAGCCAGCCGCTGAA
 CGGGTAGCGCTGGGACTCGCCCTGGAAGAAGCCCCAGCCCTTGTCTTGGCGTTGAGCTTTCTGTTGAGGTGCGCCAGCAGGT
 CTCGATGCGCTGGGACTCGCCCGCTCGCGCGTGGCGCGCTGGAGCACGGCCGCGGGTGCCTGCGGCTTCTTCTCGATCAGGTG
 GACGATGGCGTGTGATAGGCATCAGAAAGACTCCTGGTTCAGGGAAACGCTGCCCTGTCGGCTCGACGGAGCAGGGCGCGGGCA
 TTGAGCTCGCCGCGTGCCTGCTTGGCGTGGAAAGAAATGAAGATCAGAGCGCAGCCGATCTTGTGAGCTTCAAGTGGCGCGCT
 CCGGACGGTGCAGCTCCAGCCTTCTGCTTGAAGCGAGGCCACAGCCGCGAACGCTTTCCCGGGCGCGCGAAGGAACTA
 CGCATTCGCTGCTTCTCGCTGTGAGCTGCAGCGAGTTCATCTCTGCAAGGAGCGGGGAGGCTATCCAGGAGCTTCAACG
 ATCCAGGAGACGTTTACCAGCGCTGAAGGCTGCGACGATCGCTTCTCGCGCTGCAGGCCAGGTGCTCATGCGCGAACGCGAG
 CATGCGGACCTCTGTGCGCAGTTGAGTGCATACGTCGTGCTGCTACTGGTGCCTTGAACGCGACAGGTGACCGCGACCGTGT
 CCTGAGCGGCTTTCCCACTCTTCCCTTGTGCGCGAAGCGATGGCCACGACCTACACTCCAACCGCTGGTCCAGTAGCGCGA
 CTTTCCGCGAGTCAGGCGAGGGCTGTGCGCGCTGCCATCAGGTTGACCGCGAGTTCAGACAGCTTCCAGGAGCTTCAACG
 CAACTTGAGCTCAAGCATCAGCTCGACTTTCTGGAAGTGTCTGAGCGCGCGGTAGGCTGTGACGTGAGGCATCGCGCGTACCT
 CTTTGGGATTTCTACAGGACAGCGCCAGCGCTCTCGATGCGCTTTTCTGTTGGTGCCTGCTGTACCGGATACGGGTGATGT
 ATGCGTTGCTCTGCCAAACAGTCGTTCCGCAAGGCGGAGCTGCCACCTTGTGCTCCAGGTGCGAGTGCAGCTGCTGCTCT
 GAATCGTTATGGTTCAGAAACCGCGCGCATCGAACTTGAAGTGTGAGTGTGATAGGTCTGGGACGCTTCTGAGCTCGGCTCCT
 GCTGAGTGCATCACCAGCTGAACGCGCCACAGTTCGGGCAACTCCGCTCATGCCCGGATAGTTGTTGGCCAATGATGCAG
 GAACTCGCGCGCGGCTCCTTGTGCTGCAATGCGGAACACGGCGGTGAGGTGCTGGGAGTCTGAGTGCAGCGAGCAGAGGTG
 CGGCACAGGCTCACTTTCAGCGAGCAGTGAACAGGTGTCGCGGATCCATACCGACCAAGTAAGCCGAAAGCAATTGCC
 CGGAGCGCGCTGATGACCAACAGAGTCCCGCGCTGGTGTGCTGAGATTTGATGTACCCAGTCTCTCCAGCTCACGAAGCTC
 TCGCTCCATCGAACTGGGTGCGACCGTGTCTCTGTCGCGAAAGGCGCGAGATCACTTGGCTGAGGGCGTCCAGCTCGATGGG
 GTTGACCCCAACGTGATTGGCATGAACCACTGCCGAGTCGGTGAACGCGCGCGAGGCGTGGATACTCGGAAAGCGGGCTTCAGT
 GTTCATCAAAGACAGCCTCAACTGATGGACGTGCTAGCTTCCGAGCAGGGGCGAGCAACACCTCGGCTTCTCTTCTGGG
 TGTTCGCGGGGAGAACCTGGCGCACCAATGGTGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
 GGAAGTGTTCATGATGGGCTCCTTGGCAGGTCCCGCGCGCAGGTGCGCAAGCGGGTCTGGTGGCGATGAAGCGGGTGTG
 GACGGTCCGATCGGACCGTTCGATGATGCTGCGGATGATGATTTCCGCGATGCCCTTGGCCTCGGTGTTGGGGTGGTAGACCTC
 TCGCGGTACACGAACATGATCAGTCCGCGTCTGCTGATGCGCGCTCTCTCGTAGGTCCGCGCAGTTCGGTTCGCTTGT
 CGTCCGTTCTCTAGCCCGCATTTAGCTGGGATAGAGTACGACGAGTCCATCTCCTTGGCCAGCTTTCAGCGAGCG
 GGAGATTTCCGAGATCTCGGTGGCTCGATTCTCCAGGCTGGGCACTCATCAGTTGCGAGGTAGTCGACCAATATCAGCGCGG
 GTGTCCGTACTTCTTGCGCGCGCGCGCAACCTTGGCGCGCAGCTCTGTCGCGGTGAGGTGCTGCTGATGACGAGCGG
 GCTGCCATAGTCGTTTATGCGCTGGATCGCGCCAGCAGTCTGGGCAACTCTCTTCTTTCAGTGGCCCTTCATCAGCTTGCC
 CAGGTGCGCGGGAACAGGGCGGCAAGCCTGAAACAGCACTTGTGCTGCGCGCATCTCCATGCTGCGCGGCTGCGGCTGCGC
 CTGTTGGTCTGCTTGGAGCGCGGTGTGACACAGGTGAGGGCAACAGCGTTTTGCCCATCGCGGGCGGGCACCGACGATGAT
 GAGATCCGACTTCTGTAGTCCGCGCGGTGAGTGGTTCGAGATCCTTCAGGCCAGTCGGGACCCCGTTACCGTCAAGTTGTTGTT
 GAAGCGGTAATCGATGGTGTGACGATCTTCTGAGACTTGTGTTGATATCGACGAAATCGGCGTGTGGTGGTCTGGCCAAAG
 GGCGGAACGCTTCTGCTCAATCTCCTCTGAACCTCAGAGGATTTGCTGCTGGTGGTTCGAGCGGTGCGGCTGCGGCTGCC
 CGAGAGACATCAGTGCAGCGGTGTGCCGATCGCGAACGATCTGCGCGTAAGCTCGATATTGGCCACGAGGGCGGTGTTGT
 CGGCGAGCTGGCCGAGGTAGGCCAGCCCGCCAGCTTCTGGAAGGTCTTCGATCGCTTCCGACACAGTCACGACATCAACCGGAG
 CATCTTTCGCGGCCAACTCGCTGATGGCGGTGAAGATCAGCCGATGCTCATGCCGAAGAAGTCTCCTTCTGCAACTGATCGC
 CGACAAATGTCCCATGCGCGCTTGTCCAGCATCAGGCGCGGATACCCCTTGTCTCGCGCTCGAGGAGTGGCGCGGCTAAGTT
 CCAGAACACTCACTTCGAGTCTTCCCGGTGATCTGCAGAAACGTGCTTCTCTGCGCGCGGGGAGCGGATGATCTCCAGCGC
 TTTCACTTCCACCTTGGCGGACTCGATGATCACCAGACGACATGAGCCACCGCTTTCGCGCGCTCGATATCCAGGGGCTTGT
 TCTGTCAGCAACCTTTCGATGGTGGCAAAGAGGTGATTCGCGAGATCTCAATCTTGTCTTTCATGCTCGACCTCTGAATGG
 TGCGTTTGGCTGCTGATGACGCGGATTGCGCTCTTCACTCAGGTGATAGCGATGGATGGAGTTGCGGCGCATGCTTCCG
 CCCGGGTGATGAGCTCCAGGTGTGATGACGACGTTCTGCTTGTGTTGCTTGAAGCAGACGCAATGGCGGTTGGGATGG
 GGCCGAAGTGTCTTCCAGAGCAGGATGTGGATGCCCTTCCAGTCCCGGGGGGATAGCCGATATCCGAGATCTTCTGTTGCA
 GGTAGCCATCAGCGCTGACCCGCTGCTGCCGACTGGGAGCCATGTGTGGGGCTTTTGGCCCTTCTGAACTGCGTTTCTGGT
 CGCGTCTCGCGCGGCGAGCCCTTTCATGCCGAGTTCATGGGTTGTTTCCCTTCTTGAACCGAGTTGCGCTGCCAGCTCT
 GCCGATCTTCTTGGAGAACTCAGGGCTTTTTTTTTCAGCCCGAGCCGATACGCTCTGGAGCAGATCTGCTGGAGCGTTTTGTTCA
 ACCTGGCTGCCAGGACCTCGTTCCGCGATATCCGGATACAGGCGCAGCAGCAGCAGTTCCTGTTCCGCGGTCCAGGTTTCCGCGA
 TGCAGCGCTCTCTCGGTGCTGAAATTTGCGGCGGGCAGCAGCAGGCGCTCGTAGGCGATGGTTCGAGAGAGCAGAGCGAGTCA
 TTGGGAACCTTCCGTGCTGAGCGCGCCCTGAACTCGCCGCTGACACCGCGGTGAGGCGAGTGTGTTTTCGCGTGTCCAGCC
 TGCTTGCAGCGAGGCGTGGTATAGGCAAGGCGCTTCTGAAGCGCGCGCGGCTCCAGGTAATTCAACGCCGATGGTGTAGTTC
 GAGGTAGGCTTGTGCTGCTGCTGAGTGCAGTCTGTCGCGCTGCCAGCAAAATCAAGTACCAGCGAACCGACTGCGAGAAGCCA
 GCCTTCGACGAAGAGCTTGTGCGACGACGCTTGGTGCAGAGCTTGCAGCGCTTCTGTTGGGCGACATGAGCCGAGCGCGCTGC
 CACTGTTGGTGGGAGCGTAGAGTAGGCGGTGAGCGGCGAGCTCAGGGCCGATCCCTCGGCGATGAACTTGAACGTCAGCC
 CGCTGGCATCGCGTATGCCGCCAGGTAGTCGAGCCGAAGGAAGTGGCGAGATCCCGGCCAGTTCATGACGACATGGGGCAGG
 AGAGCGCTGGTTCCAGTGGGAAACGACGCGGTTTCCACCGTGTGGGCGGAGATCTCCGATCGCTGAGGTGCAACTGCCGCTT
 CAGGATCGCGGCTGGCGCAATGCGGTTTCCGCTCGTTCCGCGTGGCCCCCTTGGCATTGCGCAGGGCCATCAGCTTCTCGAC

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND
 PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September 12, 2003 Serial No.: Not Yet Assigned

Page 17 of 118

Customer No.: 21559

18/118
FIG. 2P

CTTGGCGAGGACTTTTGTGTCCATTACGCGCCCCCTTGAGATGGTTGAGCCCTGCTCGGTGGCGGAGCGGTCCAGGCGC
TCGATCTCGGCCAGCACTAGGGCGCCGGCGCGTACCAGGTTCTCGCGCGCCGACTTAGGCTTCCACCACTCTTCGTCAGGGC
CAGAACAGTGCAGAGCCATTCCGAATGCCGCCCGGCCAGTCCGGTGCCGATGCCGCCGGCATGGAGCGCGTAGCAGCGGGCCGCC
AGGGCGATCAGGCCAGCGTAGCGTCGTCGTTGGTATCGAAACCTTCGACGGTGATCTGGCGCCGGCGCTCGGCCTGTACA
TCGAGCCAGGCTGCGGAACAGGGGAGAGGTTTCAAGTTGGGTAAATTCGGTCTGTGGCCTGTCTTACCCGATCGACAATAGCC
AGGGCTGCTACGTCGCCCAGTGCGACCTCCAGGGCGGCGAGCAGGTTGGCCATCAAGAACTGCTCACGCTCGTCCGCGTGGTTT
TTTGGCTTGGGCGTCTATTCCAGTCGATCTGTGTGAATACCTCGCTGTTACCGGGTCTGGCTCGCAGTAGGACGGCCCGATG
TGAACTGGCAGTCCAGGCATTCAATGGCACAGTTCAGGCCATCGGTGTGAGCAGCTCGGGTTCCAGCATTCCGAATGTTCTGTG
CTGCCCGAGAGCGGGCAGGGCTTGAGCTTGGGGTCACTCATGGCTGCTCCCTGCCGTGCTCGGTGCTGGGCTTAGCTTGGG
AGGAGCCGGCGACGAGCAGGTCCGGCCAGTCCGGATGAGTTCCGATGGGTCCAGAAGCCGAACTGATCGCGCTCGACAGGTGGG
CCTGAATCATAGGTTTTCATGCCAGGCCTCCCACGCTCATGCCGAGCACCGCATTTGCCCTCGGCGGCCAGTTGCTGGCCTTCAGT
GCTTGATGAGTGCAGCAGGAGCATCAGCCGTTTCGATGACCTTCAAGTGTCTGCGAGGGTTGGGCGAACTCGGCCAGGTG
ATCGATGGCAGTAGCTATGTTCCAGATCGCCCTGTTTCATGTCCCGCTGGATCACTTGGATCGAACTGGCATTGGCCGCGAGT
CAGTTGCGATTTTCGTCTGCAACAGTAGGTCTGGGGGATTTGGAGCAGTGGCTTCTGCAGGGAAGTCCACAGGCACGCTGCTG
AGCGCGAACCGGAGCGACAGGCGCTGCAACATCGCAGCACCTTCAGCCTCGATTTGCAGATTCTCGGAAGCACCGGCGGCGAGC
GCCTGCTTTGGCCATGTAGCCTTGCACGTAGGCGTGAAGTGGGTCTCGGCGATCAAGCTCAGTAGGAGGTTGTTGCGGTTCTG
GAGGTTTCATATGAGGCCCTCCGAGCCGATTGCAGAGCGTCTGTCCAGAGCCGAAGGTTTCCTGCTCCTCGCTGGTCAGAGCG
ATCTGTGCGGTGCAGTCAGCGACCTTCTGCTTGATGACGCGCAGGCGCTCCTCGACTTCACTGCCATCGTCGTACGCGATTTCG
TTAGTCAGATGGCTAATGAAGTCACATCGCAGGCTCCCCACCTGACTAGCGAAATACTCGGCTTCACTGCGCCATCCGATGAGC
TGCTTTGGACGCTCCAGCCAGCGTTTGAGTTTGTAGTGGAGCTGGTGGTGGAGTACAGGTGATCGAACACCTGGCCCGGTTG
TAGATCAACCGCATCTGCAGGATTTCCAGCCGATCTTCGAGCGTATTTCCCGGAAGCGGTAGTTTCGAGGCGACCAAGTTGCTTG
GCCAGGCGCTCACGTAGCTCGCGCCGCTTCTTGGCTTTTTTCTCGTAGCGATCACGCTTTTCCTTCGCATGCTCGATCTGCTTG
TTCAACGAGTCGAGCAAGCGCGCTGCGTCTTGTAGGAGCTGCTTTTGTTCGAAGCTCAAGTAGCAACCCAGCTTCCCGAACAGG
TCGATACCGCCACCAAGTCAGCTTGGCTATTTCTCGCGTAGCGCCGGTCAATTTGGCCTGCTGGCTCCGAAGGTCGGCGGCGCTG
CCGAACCTTGTTCAGTTCCAGCACAGTTCCGCGTCGACGTCCGCGGCTTGGGTGATAGGAGTCTTCGCCATGTCAGCGCCCTCC
CTTGACCAAGTGTGCAACCGCCTCCCGCGCTCAGCGCCAGGAAGCGGTCACTCCACTCGGGAAGACCTCGATGGCCAGGTTTC
GAATGATTTCCAGCGCGGCGCTGATGTGCGATTGGAGGGTTGCCGCTTTCGAGGCGGTGCGCTGGTAGCCCGCGCGATGCTG
CGTTGCGAAACACGACGGCATCCGGGACAGTCTTTTCGAGCACAGAAATGTCCTGATGCTCATCGAAGATGGCACGCACATTCT
CGTGAATTGCCCGGAGTCAATTGGTCTGGTCCAGGCAGTTGATGACGATCTGAACATTGGGGATCCGCATGCCGAGACGCTCAT
AGGGGCGTAGGCGTTCGAGCATTTCGATGGTGCCGCGTTGAACTCAAGGCGGTAAGCATGTTGGGTTGGAGGGGGGAAACAA
CCAGGTCCGATGCAAGCACAAACATTTTCGAGCAAAGCTGAGCGCGCACCTGGGTGTCGATCAGCACCAGGTCTAGCCTTCTT
TCAGAGCGGGCATCAGGTTCCGCCAGGCGTAGCCGGCCATCGGGCGCCTGGAGCAGTAGGTTGTTGAGCTGATTGTTCTGGTCTG
TGGAAATCACGACGTCCAGATTGGGGATAATCGTCTGGAGATGATCCTCGCCGGTCCGTTATGTTGGCGGCGAGCAGGTCTG
AAATGCCGCCCTGGGCAACTTCCGGCAGCTCGTAGTACGAGGATAGGGAGGGCTGGACGGGGTCCAGATCGATGAGGAGGGTTC
GTATGCCTGCATCGGCGCAAAATGCACCGAGGTTGGCGGCGGTGGTGGATTTTCTACACCACCTTGGTGGATACAAACGAAG
TCGCTTTTCGCAATTCATGTCTACAGCCTCTAATGTGGGGTATTAGAGGCTGCTAAACAGACTCGTTGTAACGGGAGTGGAACCC
CTTCTAGAGAAACCTACGACCAAGTTGCTTTTAACCAATTGGTCTGATGTTTCGAATCCTACACGACCCACCA

FIG. 3

RL024

DNA sequence (SEQ ID NO: 3)

GTGGCGCTGACCGGTAATCCCCCTCCTGAAATTGCTGGTCGTCCCCGTCTGTGATCGGCGCCATC
CTGATCGGCGTGAGCATGATGGGCAAGAAAGAAAGTGCGCAGTCACAAGGCGCCGCAACCCCG
ACGGTAACGTTCGGAAGAAGCGGCAACCCCTGGGCATCGACGGCGACACGCCCCGCCGACACACTA
CGCACCATCGTGGCGGAAAGCCGGCAGCTCAAGGACCAGATCAGCAAGGTGATCCAGGAGAAT
GACTCGCTCAAAGCCGCCAATGAGAACCCTGCAGGGCCGCCTGCGCAACATCGATCAGAACATC
GAGCAGAAGCTCAACAACACCGCCCAGGAACTGCAGCAACAGCAGGAAAACCGTAGCCAGACG
ATCCTGGACCAGGTACAGAAACGGCTCGAGAACCCTAACCCACATTCCCGAGGCCGGTGACACC
GACCTGCCCCGTAGGATTCGGCGTGCGACCAAGGATGGCCAGCACTTTCAGGGAGCGGGCTCG
TCTTCATCGGATATCGTCTGGATCGAGCCCCAGGACGCCCCGCGCGGTTGATGCCAATGGCCAG
CCGCTGGCCGCCGGCTCCACCACCCAACCGAGCGGATTTCAGCTTCCCGACCTCCTTCGGCAAT
GCGGTTCGATCGCGGACAGAACGCGCTGGAGCGGATCGATGACGGGCTGCACCCCGTCGGCCAA
CAGCGATCTGACCTGGAACCGCAAGCTCGTCCGTAAGACCTACACGCTGCCCGAGAACTCG
ACGCTCATGGGCTCGGTGGCCATGTTTTCGCTGATCGGTCTGTGTGCCGGTTCGACGGGACGGTC
AATGATCCTTACCCGTTCAAAATCCTCATCGGCCCGGACAACCTCACCGCCAACGGCATCGAG
CTGCCGGACGTTCGCCGGCGCGGTAGCCAGCGGGACCGCCTCGGGCGACTGGACACTCTCCTGC
GTGCGTGGGCAGATCCGCAGCCTCACGTTTCGTGTTCAACGACGGGACCGTGCGCACCTTCCCG
GCGCCGGCCGAGGAGGTGAATGACAACCAGAGCAACAACAACAGACCGCCAGCGCCGACCAG
AAAACCATCCAGGGCGGCCTCGGCTGGATCAGCGACCCCTACGGCATCCCATGCATCGCCGGT
GATCGCCGATCCAATGCCAAGGAGTACCTGGGCAATCAGAGCCTACTCACGGCTGCCGGGGCC
GGCATTGCCAAGCTCCTGGACGCCGACGAGAACAACACCAAGTACCGTCTTCAGCGGCAACGGC
ACCAGCTTCGGGACGACCGGAACCAACAGCAACTCGGCCCTCAACAGCATCCTCTCCGGCGGC
GTCAGCGACATCCGGCAGTGGATGAACAAGTTGTACGGGGAGGCCTTCGCCCGCGTCTACGTG
CAGCCAGGTGCGCGGGTTCGAGTGCATCTCGATCAGCAACTGGCGATCGACTATGAACTCAAG
GGCCGCAAGGTCGATTACAGCTCTGGAGCCGCTCATGCAACAGCAGACTTGGACTAA

Protein sequence (SEQ ID NO: 127)

VALTGNPLLKLLVVPVIGAILIGVSMMGKKESAQSQGAATPTVTSEEAATLGIDGDTPADTL
RTIVAESRQLKDQISKVIQENDSLKAAENLQGRLRNIDQNIQKLNNTAQELQQQQENRSQT
ILDQVQKRLENLTHIPEAGDIDLPGVGFVRPKDQHFQAGSSSSDIVWIEPQDARAVDANGQ
PLAAGSTTQPSGFSFPTSFGNAVDRGQNALERIDDGLHPVGQQRSDLENRKLVRKTYTLPQNS
TLMGSVAMFALIGRVPVDGTVNDPYPFKILIGPDNLTANGIELPDVAGAVASGTASGDWTLSC
VRGQIRSLTFVFNDDGTVRTFFPAPAEVNDNQSNNNQTASADQKTIQGGLGWISDPYGIPIAG
DRRSNAKEYLGNQSLTAAGAGIAKLLDADENNTSTVFSNGTSFGTTGTNSNSALNSILSGG
VSDIRQWMNKLYGEAFAAVYVQPGARVAVHLDQQLAIDYELKGRKVDYSSGAHATADLD.

FIG. 4

RL025

DNA sequence (SEQ ID NO: 4)

ATGATCCGGAAGTCGACAGGCTCGCTCTTGCTAATGCTTGCCCTACCCACACTGGCCACGCG
GTGGAGATTCTGCGCTGGGAGCGCATTCGGTTGGCCATTCCATTGACGGTCGGCCAGGAACGC
ATTGTTTTTCGTCGACAGAAACGTGCGAGTTGGGGTTCCTCGGGATCTGCAGGGCAAGCTGCGC
GTCCAGAGTACCGGCGGCGCACTCTACCTGCTCGCCAACGAGCCGATTCTCCAGCGCGCCTG
CGCCTACAGGACGCGACCAATGGCGAGCAGATGCTCATCGATATCGCCGCCACCGAAGCAACG
GCCGACCAACAGCCGCGCGAGCCGGTCAGGATCGTCGCCGCGAGCCAGTGGATCCGCATTAT
GGCCAGTCCCGGGAAGCCAGCCATCGGCAGCAGCGAAACAGACCGAGCACGCAGAAGCACCG
AAGGCCGTGCCGCGCGAAACGCCCCGTCCCCGTGGTTCTGACGCGCTATGCGGCGCAGATGCTC
TATGCCCCGCTTCGCACGGTGGAACCGGTGGATGGCGTCGGTCAGGTGCGCGTCAAGCGACAG
CTCGACCTGACCACCCTGCTCCCCAGCCTACCCATCACGGCTACCGCCTTGGGCGCCTGGCGG
CTGGACGACTACTACATCACGGCGGTGAAGCTGCAGAACGCCAGCGCCCAGCACCTGGCCCTG
GATCCCAGGGACCTGATGGGCAATTTTCGTCGCCGCGACCTTCCAGCACCCGTACTTGGGGCCC
CGGGGCGACGCCTCCGACACCACTACCGTGTATCTGGTGACGCGCGGCCGCGGCCTTGCCGAC
GCGCTCCTGCCCTCTCCATCAGCCAGATCGATCCCAAAGGAGGCCGTGCTGGCGCTGACCGG
TAA

Protein sequence (SEQ ID NO: 128)

MIRKSTGSLLLMLALPTLAHAVEILRWERIPLAIPLTVGQERIVFVDRNVRVGVPRDLQGKLR
VQSTGGALYLLANEPIPPARLRLQDATNGEQMLIDIAATEATADQQPREPVRIVAGEPVDPHY
GQSREAQPSAAAKQTEHAEAPKAVPRETPVPVVLTRYAAQMLYAPLRTVEPVDGVGQVRVKRQ
LDLTLLPSLPITATALGAWRLLDDYYITAVKLQNASAQHLALDPRDLMGNFVAATFQHPYLGP
RGDASDTTTVYLVTRGRGLADALLPSSISQIDPKGRRGADR.

FIG. 5

RL026 : DNA sequence (SEQ ID NO:5)

ATGAGTTTCAGAAAACACACTGCGCAACAGCAGGCACACATCAACACCTTCCGGTTCATCACC
GGCTTCCTGTGCATGGTCATCGTTGTGCTGGCCTACTGCGTCTGGGAAGCCCGTAAGGACCTC
TGGATCCACATTCCGCCCCGACTTGCGCTCAGGAAGCACCCGGTTGTGGTGGGACATTCCGCCA
GAGAGCGTCTATGCGTTCGGCCTCTACATCTTCCAGCAGGTGCAGCGTTGGCCCAAGGACGGC
GAGGTGGACTACAAGGGAAACCTGTTCCGCTACGCTGCCTACCTCACTCCCTCCTGCAAAGTC
TTCCTGGAGAAAGACTTTGAGTTTCGTCGTAACGCCGGCGAGCTCAGGGGTCGCGAGCGCACC
ACCTCGGAAATCCCCGGTCGAGGCATTGGCGAGAGCAATGGCCGCGTGATCCAGCACTCGATC
AATGACTGGACCGTCAACTTGACATGGACAGCACGGAGTATTACGCCGGCGAGAAGATCAAG
CGGGCGCTGGCCCCGCTACCCGTTGCACGTTATCCGCGCCGACGTCGACCCGGAAACCAATCCC
TTCGGCCTGCAGTGGGACTGCTACTCCGACACGCCTCAACGTATCGAGCTTGAGGAGCCGGCC
GCCCCACCAAGCGGGAGGGAGGTCTATGA

Protein sequence (SEQ ID NO: 129)

MSFRKHTAQQAHHINTFRFITGFLCMVIVVLAAYCVWEARKDLWIHIPDLRSGSTRLWWDIPP
ESVYAFGLYIFQQVQRWPKDGEVDYKGNLFRYAAYLTPSCKVFLEKDFEFRNAGELRGRERT
TSEIPGRGIGESNGRVIQHSINDWTVNLDMDSTEYYAGEKIKRALARYPLHVIRADVDPETNP
FGLQWDCYS DTPQRIELEPPAAPT KREGGL.

FIG. 6

RL027

DNA sequence: (SEQ ID NO: 6)

ATGCCCCGAAGAACATCTGTTTCAGGATGGAACCCCTCAGCTTCCTGCCGACCCGTTTGAACCGG
CAACCGGTAGTCATCGGCGGCCTGACCGCAGACGAAATGTGGATCACGGTCTTCACCAGCGGA
GCAGCCGGGTTCGTTCTTGGCATCCCGGCTGCCTTGGTCGCAGGTAACGCTGCCTGCATTCCA
CTGGGCGCGCTGCTGGTCGGCGCCCTCGGCCTAGGTATCGGCAGCCGCGTCCTGCGGCGGATG
AAGCGGGGGCGGCCCCGATACCTGGTTCTACCGCCAGGTGGAGATGGCCCTCTCGCTGCGCTTT
CCCGTCTTCGGCAACCGTCGCCTGGTTACGCGCTCCGGCGCCTGGACCAGTCGACGCACGGAG
TCCCCATGA

Protein sequence: (SEQ ID NO: 130)

MPEEHLFQDGTL SFLPTRLNRQPVVIGGLTADEMWITVFTSGAAGFVLGIPAALVAGNAACIP
LGALLVGALGLGIGSRVLRMRGRPDWTFYRQVEMALSLRFPVFGNRRLLVTRSGAWTSRRT
E SP.

FIG. 7

RL028

DNA sequence: (SEQ ID NO: 7)

ATGCTGAAACTCACCTCCAGAACTGTCCGCCCTCTGCCAGAGCCTGGCCGCCATCACTTTG
GCGCTCCCCGGTATCGCCTTGGCTGCACTCCCCAAACCCGAGGCACCTAGCCGTGGGGAGGGA
TCGGGCATCATGCAAACCATCCAGAACTTCGGCTATGACGGAGCGATGCTCCTCGCGCTGCTC
ATCTGCGCGGCTGTCTTTCTGGGGGTCGCTTGGCATACTACGGCACCTATCACGCCATCCAT
GACGGGAAGAAGAAGTGGTCGGATCTCGGAGCGGGCGTAGCCGTAGGTGTCGGCCTGCTGATC
TTGATCATTTATCTCGTCACCAAAGCCACCGCCATCATGTAA

Protein sequence: (SEQ ID NO: 131)

MLKLTLQKLSALCQSLAAITLALPGIALAALPKPEAPSRGEGSGIMQTIQNFGYDGAMLLALL
ICAAVFLGVAWHTYGTYHAIHDGKKKWSDLGAGVAVGVGLLILIIYLVTKATAIM

FIG. 8

RL029

DNA sequence: (SEQ ID NO: 8)

ATGAGCATGAGCGGAGCCCAGACATCAGCGTTCCAGGCCGCCGCTGGCTTTCCCCCATCGGCC
GGCGAGGGACTGTTTCATTGGAGCAGCGATGACCTTCCTTCTGCTGTGGTCCGCCTGGGCGATG
TACAGCACCTGGCGCGGCTGGGCCACCAACAACCTTCGACAGCGCCACCGGTGGCGCTTCCGC
GATCCCGGATCTTGGTCCTCCTCGGCATCACCTCTTTCTTCCTCCTCAGCTGACCCATACGGA
GACACTCATGCTGAAACTCACCTCCAGAACTGTCCGCCCTCTGCCAGAGCCTGGCCGCCAT
CACTTTGGCGCTCCCCGGTATCGCCTTGGCTGCACTCCCCAAACCCGAGGCACCTAG

Protein sequence: (SEQ ID NO: 132)

MSMSGAQTSFAFQAAAGFPPSAGEGLFIGAAMTFLLLWSAWAMYSTWRGWATNNLRQRHRWRFR
DPGSWSSSASPLSSSSADPYGDTHAETHPPETVRPLPEPGRHHFGAPRYRLGCTPQTRGT.

FIG. 9

RL030

DNA sequence: (SEQ ID NO: 9)

CTGATCTGCACGAGATTGCGCGTGAACACTCCACATCCATCCCTTCGCCGAAGCTGCCTGGCC
GTCTTGGCCTGCAGTGCCTGGTCGCACAGGGAGCTTTCGCAGCGAGCGCCTCCGAGCAGGCG
AACCTGGAGGTGATGATCCGGCAGCTCAACGCCCTCGAGGACACCGCCCGCCGAGTGCCCAG
GGCGCCGATGAGCCCGGACAGCGCTTCTACTTCGACTACCCGCGCCTGGCCGCTGACCTGCAG
CGCATCCGCCAAGGCCTGCAGGACTACATGACGCCCGAGCCGCGCCCAACCGCGTGACCCTTCC
GACTTATCAGGGAATTACACCCTGCGCGGAGGGCCGATGCCATGA

Protein sequence (SEQ ID NO: 133)

LICTRFAVNTPHPSLRRSCLAVLAC SALVAQGAF AASASEQANLEVMIRQLNALEDTARRSAQ
GADEPGQRFYFDYPRLAADLQRI RQGLQDYMTPSRAQPRDPSDLSGNYTLRGGPMP

FIG. 10

RL031

DNA sequence: (SEQ ID NO: 10)

ATGAGCATAAAACAGCCCTTCGAATACCATGTCGAGAACATCGTCATTCCCTACAAAACCTC
ACCAAGGGCGTCGCGATGTTCAAACACAAAGAAGACACCTTGGAACCCGACGACCACGCCTTG
CTCAACCCCTCTGCGCTGGGCCGAGGTCGTGCGTCTGGGCCAGGAAGGCTGGGAGCTGGTGAGC
GTTCAAGCCACTCATGCGGGGCGTAACCGAGATCGGTAATCAAAACGCCCAAGGCTGGGCTTGG
GGCGTCGCTCTGCCCCTCAGCTACCTGCTGTTTTTCAAGCGCGCAACCTCATAA

Protein sequence: (SEQ ID NO: 134)

MSIKQPFYHVENIVIPYKTLTKGVAMFKHKEDTLEPDDHALLNPLRWAEVVRLGQEGWELVS
VQPLMRGVTEIGNQNAQGWAWGVALPVSYLLFFKRATS.

FIG. 11

RL032

DNA sequence: (SEQ ID NO: 11)

ATGCTTAGAAACATCTCTATTGGAGTTTTGCTAGCCATGGCTGCTATGTTGGGCAGTTATGGG
GTGGCTGCCGCTACATTACGATGCGGGTCGGCAATTGTTAGTGAGGGCGACTTGATTGATGAT
GTGCTTAGAAAGTGCGGCAACCCTGATAGCCGTAAAATTGAAGGGCCCGCAGTGGATGGTAGT
GGCTATATAGTGCGGGGGGCTGCTACTGTCGAAAACCTGGGTATATGGACCAAGGAATGGATGG
TACCAGAAGCTTAGGTTTGTTCGATGGAAGACTAGTTCAGATAAAAGGCAGTATGGACTAG

Protein sequence: (SEQ ID NO: 135)

MLRNISIGVLLAMAAMLGSYGVAATLRCGSAIVSEGLIDDLVLRKCGNPDSRKIEGPAVDGS
GYIVRGAATVENWVYGPRNGWYQKLRFVDGRLVQIKGSMD

FIG. 12

RLO33

DNA sequence: (SEQ ID NO: 12)

ATGAACTTATCCTTGATTTCGACGGACGCCTTCTAAATCCAAGCAACATGCTAGAGGCCCTA
TCAAAGCAGGAAAAAATACAAGCATCAGCATAAGCAACGCGCAAGCATTAAATATAGAACT
CTTCTCAAGGCAACAACCACTGCAGAAAACACAAAAACCTCTCAACAACCTTCAACGGCGCA
GAGCTGACTGCTAACAACCTTCAGCAAGTCATAAACTCAGCAGGATCACTAACCAGAGTATCC
ACAATAGCCGCACAAGCCATTAATATAAACACACTTCTTCCGCAATATCTACAGCAGGCAAC
TCAAAGAATTTTAGCGCAGAATTCAATGGAGCCCAACTCAGCAGCGACAACCTACTTAGAGCA
GTAAATGCGGCAGGAACAAACACCAGCATAAGCGTCAATACCGCACAAGCGGCAAATATAACC
GCCCTTCTTCAAACATTCATGCAGCAGGTGACACAAAAACATTCAGCGCAGAGTTCAATGGC
GCTCAACTTACTTCAAACAACATTCAACAAGCTTTAGACGCCGCGAGGAACCCGAACATCCATT
AGCGTCAATACCGCACAGGCGGTTAATATAAGCACCCCTACTAGCCCTCATCAACTCTGCCAAA
GACACGAAAAAGTTTAGCGCCGACTTCAATGGTGCACAACCTAACAGCAGACAACCTTCAGCAA
GCGATCAGCGCTGCGGCCTCGGGTACCAATATCAGCGTCAACACCGCTCAGGCGGCGAATATA
TCCACCCTTTTACAGGCCATCAACATCGCGGGCAACACTAAAAAATTCAGCGCCAACCTTTAAT
GGTGCCCAACTCACTTCAAACAACATCCAGCAGGCGCTCCGAGCGACAGGATCAAACACATCA
ATCAGCATGAACTCCGCACAATCCGCCAACCAAGCACTCTACTTGAACCTCTAGACATAGCA
AGTTCCAGCAAGCAATTCCAAGCCAATTACAACGGTGGCATGTCTAATCCGAACAACCTACAA
CAGATAGTTTTCCCGTGCAGGCGCCAGTACAACCGTGTTATTTCCGACGCACAAGGCCTACC
AATCGCAAATATCCTTACCCTTATATCATCTGCCGGATGAGACTTATAGCCGTGGATGAAAC
ACACCATCCACGGCTATACCCTAG

Protein sequence: (SEQ ID NO: 136)

MKLILDFDGRLLNPSNMLEALSKAGKNTSISISNAQALNIETLLKATTTAENTKNLSTTFNGA
ELTANNLQQVINSAGSLTRVSTIAAQAININTLLSAISTAGNSKNFSAEFNGAQLSSDNLRA
VNAAGTNTSISVNTAQANITALLQTIHAAGDTKTFSAEFNGAQLTSNNIQQALDAAGTRTSI
SVNTAQAVNISTLLALINSAKDTKKFSADFNQAQLTADNLQQAISAAASGTNISVNTAQANI
STLLQAINIAGNTKKFSANFNQAQLTSNNIQQALRATGSNTSISMNSAQSANQSTLLELLDIA
SSSKQ
FQANYNGMSNPNNLQQIVFPCRRQYNRVYFRRTRPTNRKYPYPYIICRMRLIAVDENTPSTA
IP.

FIG. 13

RLO34

DNA sequence: (SEQ ID NO: 13)

GTGCAGTGGACTCACGAACAGTCACCGATCATCCAGTCTGAAGGCACCGAAGATCCTGGTGCGA
GCCTTCGCAGGCACTGGCAAACTACCACCCTGGTGGGCTTTGCCAGGTCTGAACCTTACCCTG
AGAATCCTCTATCTCTGCTACAACAGCTCGGTGGAGAAAGCCGCGAAGGGCAAGTTTCCCCGC
AACGTAGTGTGCAAGACCGCCACAGTCTGGCTCATGCGGTGTACGGCATTAGTACGCCCAC
AAGAAGACGAAGAACCTGCGACTGACCGATATCGCCCGCGGACTCGATACCCAAGACTGGGAG
TTGGTACGTGACGTGCTGGCCACGCTGAACAACATACATGGCCAGCGCCGACGCGGAACCTCGGC
CGACCGCACTTCCCGCGCTTCCGCGACAAGGCGTTCCTCACCAGTGTCTAGGAACGCTTCCTC
AAGCAGGGCCTGGACATGGCGCGAGTAGTCTGGAGGCGCATGGTCTGATCTCCAGGACACCGGC
ATGCTGATGCCCCCTTGACGGCTACCTGAAGCTGTATCAACTGAGCAAGCCCCGATTTGAGCCAG
CGCTTCGACTGCATGCTCCTGGACGAGGGGCGAGGACATCAACCCAGTGATCGCGGACATTGCC
CATTGGCAGCGCATCAGAATGGCTATCGTTCGGCGATCCCCATCAGCAGCTCTACCGGTTTCAGG
GGCGCAGAAGATGCCCTGAACAGCGACTGGATGGCCGGCGCCGAGGAGCACTACCTGACCCAG
AGCTGGCGATTTCGGCCCCGCGATCGCACACGTGGCCAAACATCATCTCTCTCTACAAGGGCGAA
ACACGGAAACTTCAAGGACTGGGTCCGCGAGACGCTGGTGAAAAAGTCCCTCCCGCCGGACCTT
CCTCACCGCACTTTCATTACCCGACCGTTATCGGCGTCATCGAGAATGCCCTGCAGCTGGTC
CGCAATCATCCGGAGCCCAAATTCCTACTGGGTAGGCGGTATCGACAGTTACTCGCTGCGCGAC
CTGGAGGATCTGTACGCATTACAGCCGAGGCCTGCGCCAAAACGTCCAGAACAGAACTGCTC
CGTGACTACCGCGACTACACCCAGTACGTGGAGATCGCCGAGATCAGCCAGGACGGTGAGATG
CTTCGCTCGATCAAGATCATATCGACCTACCCTGATCTGCCCTGCGCGGATCCTTGAGCTTCGC
TCACTGACCCCTTGACGATGAGCTGGACGCAACAATCACCTGACCACCGCACACAAGGCCAAG
GGGCTGGAATGGGATTTGTTTGCCTGTACGACGACTTCAACGCGGACCCGCTGGCCCCCGAC
ACCGACCCAGGCAGCGCGACGATGAGTTGAACCTGATCTACGTCGCAGTGACCCGCGCGATG
AAGATCCTTGCCATCAACAGCCTGGTGCTGTGATCATGCAGCGGTACGTGGACGACAGAAAA
CTGAAGGAGCAGATAGCTAGCTGTAAAAAATGA

Protein sequence: (SEQ ID NO: 137)

VQWTHEQSPIIQSKAPKILVRAFAGTGKTTTLVGFARSNPTLRILYLCYNSSVEKAAKGKFP
NVVCKTAHSLAHAVYGIQYAHKKTKNLRLTDIARGLDTQDWELVRDVLATLNNYMASADAELG
RPHFPRFRDKAFLTSAQERFLKQGLDMARVVWRMVDLQDTGMLMPLDGYLKLYQLSKPDLSQ
RFDCLLDDEGQDINPVIADIAHWQIRMAIVGDPHQQLYRFRGAEDALNSDWMAGAEHHYLTQ
SWRFGPAIAHVANIILSYKGETRKLQGLGPQTLVKKSLPPDLPHRTFIHRTVIGVIEINALQLV
RNHPEPKFHWVGGIDSYSLRDLEDLYAFSRGLRQNVQNKLLRDYRDYTOYVEIAEISQDGEM
LRSIKIISTYPLPARILELRSLTLDDELDTITLTTHAKAGLEWDFVCLYDDFDNADPLAPD
TDPGKRDELNLIYVAVTRAMKILAINSLVLSIMQRYVDDRKLKEQIASCKK.

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND
PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September 12, 2003 Serial No.: Not Yet Assigned

Page 29 of 118

Customer No.: 21559

FIG. 14

RLO35

DNA sequence: (SEQ ID NO: 14)

ATGTTTCGGGTCGCTGATCGGCGCAATCATCGTGGAGTGGGTATGCCTGTATTTCTTCTGGCCT
GACGCGGGCTGGAAGCATGCCCAGGCCATGTTTGAGTACGAACTCAGTTGGCTGTTCGAGGGG
CTGCTACACAGCGTCGTTCGTGCAGGAGCCAGGTTCGAACCGCCACCTGGCTGGCCCAGTTGGCC
TATGACTGGTTGTTTCGTGAAGACCGGGATGGTCGACTGGATGACCAACATGACTACCATCGCG
CAGGCCCCGGCCACGGAGCCCCGCTGGACGTTTCGCTATCTCACCGCCCACGGTGTCTCCACGCTG
CAGAACTACGGCCTGGCCGCGCTGTACACGGTGCTGACATTCGTCGTGCGCCTGGTGATCCTG
GTCATGACGATCCCGTTATTTCGTGATGGCCGCGTTCAACGGCCTGGTGGACGGCCTGGTGCGC
CGGGACCTGCGCAAGTTCGGCGCCGGCCGGGAGTCCAGCTACCTCTACCACAAGGCGCGCGGC
AGCATCATTCCGCTAGCGGTTCGTCCTTGGACGCTCTACCTGGCAATCCCCATCAGCATCAAT
CCCCTGCTCATCCTGTTGCCCTGCGCCGCGCTGCTCGGCGTAGCGGTATGCATCACAGCATCC
ACCTTCAAAAAGTACCTGTAG

Protein sequence: (SEQ ID NO: 138)

MFGSLIGAIIVEWVCLYFFWPDAGWKHAQAMFEYELSWLSQGLLHSV VVQEPGRTATWLAQLA
YDWLFVKTMVDWMTNMTTIAQARPRSPLDVRYLTAHGVSTLQNYGLAALYTVLTFVVRVLVIL
VMTIPLFVMAAFTGLVDGLVRRDLRKFGAGRESSYLYHKARCSIIPLAVVPWTLYLAIPISIN
PLLLLP CAALLGVAVCITASTFKKYL.

31/118
FIG. 15

RL036

DNA sequence: (SEQ ID NO: 15)

ATGAAGTTGAAGAATTTCTTACAGCCTTTTGTATAGCGGTTTCTCCACTCCGAGTGCTGCGCTCAAGCTGCTCCGCA
TGCTCGGTGGCGCCTTGATGTTGTGCGTGCTATGCAGCCTGATATTCAGTGTGAGCATGGTTTTAAACCATCAGGT
GTCCCTCAGTCGGCAAGCTATGAATGTGGCTATGTACGAAGCGCAGCTTTATTTTCGAGCAGCGCGAGGCGTTGCTC
AATCACTTGAGCGGCAATGTCGTGCCCTTGGCCGCGGGTAGAGCGCTCGTCAACGAAGCGCCGAACAATGTGAGCA
TCCTGCCGTTGAGTGACGAGGGCGAGGTCTGCTATAACCGCTCGCACGCTCGGTGATCTCCGGGAAAAGCGGCT
GGCACTGATGTATCTGGTCGATACCGACAAAGGCCCTCTGGTTTACCGGCTTACCGCGGATGGTAGGCCCTCGGCA
GCGATATCCAGCACGATAACCAAAGAGGTGTACCGAGCCTTGCTGGCGACTCCGTGCGCGCCTGTTCACTGGGTGA
CTGACGGTGGTACCCCTCAACGGCTGTACCTTTTTGAATCCTTAGGCGATGAGCCGGGCGAGGGGTGGCTAGGCCT
GGAGATTCTCGGCGAAGACCTCGATTGATGTTGCGCCGGAATGATGCCGAAACTACATGCTGCTGGATCAGCAT
GGGCGAGTCTACTCGCTACGGACGCAGAGGCGCTGGGGAGCGGTGCGTCGCGGACGCTTTTGGCTGGAGACGGCT
TCGGTTTCATCGGTGCTGGCCACTGCGCGCAGCATATGGTGCTTTTCCAGCACGTGGGGTCTTCGAGCTGGGATCT
GATCTATCACATCGGTATCGGTGCGCTGTTGCTGGCTCTGTGGCTCCCTCTGTTACTTGCCTCTGCGTTGGCACTC
GCAGTCGGCATCCTACTGCATTGGCTGGTGCGGAGCATCGAGCGACGCTTGATAGAGCCGCAAAGCGACGCTTG
AAGCATTGAAGGAGAGCGAAGCCTTTTCCCGTGCACTTATCCAGGCCGCGCCGTCGCGCTGTGCGTGCTGCGTCG
TGCCGACGCCGAGTGGTCTCGGAAAATCCCCAGGCGCGCCAATGGCTGGGTGATAGCGAGGCGATTGCCACGAC
GCGCCGAGATGGATTTCCAGGCGTTCGAGGAGGTGTGAAGTGTCTGGAGAAGAACTGGAACCGAGGCGAGGCG
TACATCTTCATCTCAATTACACGCCCCACCGCTATAACCGGTGAAGACGTATGTTCTGCGCCTTCAGTGAAATCAG
TGACGCAAGCGGATGGAGGCGGAACTGGCTCGCGCAAAATCCCTGGCGGATGCTGCCAATGAAGCCAAGACGCTG
TTTCTCGCCACCATGAGCCATGAAATCCGCACACCTCTGTACGGCATGCTTGGCACGCTTGAGCTGCTTGGGCGTA
CCGAGCTGAGTCGGCAGCAGGCCGCTTACCTAAAGGCAATCCAGCATTCCTCGTCGACCCCTGCTGCAACTGATCAG
CGATGTGCTTGACGTATCCAAGATAGAGGCCGGCCAACCTGGACCTAGAGTGCGTGGAATTCTCCCCGCTGGAATTG
ACCGAAGAGGTCGTGCACTCGTTACCGGTGCGCGCGAGGCCAAGGGGCTGCAGTTGTATACCTGCCTCTCTGCGG
AGCTGCCGCTGCGCATGCGGGGGCGCGCGCTGATTCGGCAGATTCTCAACAACCTGCTGAGCAACGCGGTGAA
GTTACCGACAATGGCTATGTCAACGTCCACCTGAAGGCCAGCGTGGTCGATGCCGAATGTGTGATGCTGACCTGG
CAGGTCAACGATACCGGCATGGGGATCAACGTGAGGATCAGCCGCTCTGTTCAACCGTTCTACCAGATACGCC
GCTCCGAGCATCCGTCGAGGCACGGGCTCGGCTTGTCGATCAGCCAGCGCTGGCGCAGCTAATGAATGGCAG
TCTGAAACTGGTCAGTGAGCTGGGGTTGGGCAGCAGCTTTAGCCTCAGGCTTCCGCTTGAGCGGATCGCGATGCA
GCTGAGCCGAGGACCTAGCCGGGTGCGCGCTCAAGTGTGCGCGCTGTCGCGACCTAACGGAATGCCTGTGTG
GCTGGATCTCCGCTGCGGTGGAAGGCCATGGTCGCGACGCCGAGGTGCTGGACGAGGCGGACGCGACCTCGCT
GCTGGTCAAAGTGTACTGCTGAGGGGGCGCGGATGTTTCAAGCATGGCCAGGATGCCGGGTGGAGCTTTCCCT
CAGGGTGATATGGAGCCGAGGCACAGGGCCGCGACTGGCTGCTCGGGCTCAACAACCTGAACGGCCTGCATCGTG
CTCTGGGCTGGCCATGGCGCTCTGCTGATCCTTCGACGCCGCGATACGGCTGGCTCCGTTGCGCAATCTAGG
TCTCCGCTCCTAGTGGTGAGGATAACCGCATCAACAGTTGATCTTGAGGGACAGATGGAAGCGCTGGGCTGC
AGCGTGAGCTGCTCTTCGATGGTCGCGAGGCGTGTGCTGCATGCCAGACGGCCTGCTTCGACGTGGTGCTCACCG
ATATCAACATGCGCAACATGAACGGATACGAGCTAACCGCGGAGCTACGGCGCCAAGGGTTCCGGCAGCCGATCAT
CGGCGCGAGCGTGAACGCCATGCGTGAGGAGCGCGAGCGCTGCATGTCCGCCGGGATGAACGATTGCTGGTCAA
CCGGTGGATCTGAATGCCCTTCAGAACTGCTTGATTAATATTCTCAAGGTGGATCGATGA

Protein sequence: (SEQ ID NO: 139)

MKLKNFLQPFDSGFSTPSAALKLLRMLGALMLCVLCSLIFSVSMVLNHQVSLSRQAMNVAMYEAQLYFEQREALL
NHLSGNVVPLAAGRALVNEAPNNVSIPLSDGGRGLLLTARTLGDLREKRLALMYLVDTDKGPLVYRLTADGRPSA
AISSTITKEVYRALLATPSAPVHWVTDGGTPQRLYLFEISLGDPEGEGWLGLEILGEDLDSMLRRNDAGNYMLLDQH
GQVVLATDAEALGSGASRTLLRGDGFFIGAGPLPQHMLVLFQHVGSWDLIYHIGIGRLLALLWLPLLLASALAL
AVGILLHWLVRISIERRLIEPAKRRLEALKESEAFSRAVIAQAPVALCVLRRADAADVLENPQARQWLGDSEAIAND
APRWISQAFAGVKCSGEELETEAGLHLHLNYTPTRYNGDVLFCFSEISARKRMEAEARAKSLADAANEAKTL
FLATMSHEIRTPLYGMLGTLELLGRTELSRQQAGYLKAIQHSSSTLLQLISDVLVDVSKIEAGQLDLECVEFSPLEL
TEEVVQSFTGAAQAKGLQLYTCLSAELPLMRGAAASIRQILNNLLSNAVKFTDNGYVNVHLKASVVDACEVMLTW
QVNDTGMGINVEDQPRLFEPFYQIRRSEHPVAGTGLGLSISQRLAQLMNGSLKLVSELGLGSSFSRLPLERIAMQ
AEPQDLAGCAVQVLAPVRDLTECLCGWISRWGGGRAMVATPRSLEADATSLLVKVLVLLLEGAPMFEAWPGCRVELSP
QGDMEPQAQGRDWLLGLNNLNLGLHRLGLAHGRLADPSTPPIRLAPLRNLGLRVLVVEDNAINQLILRDQMEALGC
SVELFDGREALLHCQTACFDVVLTDINMPNMNGYELTAELRRQGFQPIIGATVNAMREERERCMSAGMNDCLVK
PVDLNLQNCLINILKVDK.

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND
PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September 12, 2003 Serial No.: Not Yet Assigned

Page 31 of 118

Customer No.: 21559

FIG. 16

RL037

DNA sequence: (SEQ ID NO: 16)

ATGAGCTGGAAATCCTATCGGGTGCTGGTGGTTCGAAGATCAGCCGTTTCAGCGCGAATACCTG
CTCAACCTGTTTCGCGAGCGCGGCGTGACGTACCTGGTAGGTGCCGGCGACGGCGCGGAGGCG
TTGCGCTGCCTGAAGCAGGACAGGTTTCGACCTGATCCTCAGCGATCTGATGATGCCGGGCATG
GATGGTATCCAAATGATCCTGCAACTGCCGTATCTCAAGCATCGTCCGAAGCTGGCGCTGATG
AGCTCCTCGTCGCAGCGGATGATGCTCAGTGCCAGCCGGGTCGCCCAGAGTCTCGGCTTGTCG
GTAATCGACCTGTTGCCCAAGCCGACTCTGCCCAAGGCCATCGGCCAACTTCTGGAACACCTG
GAAAGATGCCTCAGGCAGAAGCTGGAGCCGGAACCGACGAGACTCCGCATGGGCGCACGGCG
TTGCTGGATGCCCTGCATAACGAGCAACTGGTGACCTGGTTCCAGGCTAAGAAATCCCTCCAC
ACCGGGCGCATAGTCGGCGCCGAGGCGTTGATACGCTGGAGCCACCCGCAGCATGGCCTGTTG
CTGCCCAGCTGTTTCATGAGTGATGTGACGCTACCGGTCTGCACGAGGCGTTGCTCTGGCGC
GTGCTCGAACAGACCCTGAACGCCCAGGAATCGTGCGCGCAGGGCGGGTTACGAGATTCCGGTT
TCGGTGAATCTGCCGCCGCACCTGCTCGATAACCAGGAACCTCCGGATCGACTCTATGAGTAC
GTCCGCGCTCGCGGGGCTTGTACCAGCTCACTATGTTTCGAGTTGACCGAGAGCAGTGTCACA
ACTCTGTCAAGTAACTACTATGCAGGTGCCTGTGCTTGCATGAAAGGGTTCGGATTGGCC
CAGGACGACTTTGGCCAGGGTTACAGCTCGTTCTATAACCTGGTTCACGACGCCTTTACGGAG
CTGAAGATCGACCGCTCCCTAGTCCAGGGATGCGTAGAGGATAACGGCCTCAATGCAGCTGTC
ATCAGTTGTATTGAGTTGGGTCAACGCCTGAATCTCGACGTGGTGGCCGAAGGCGTGAGACC
TGCGAGGAAGTGAATCTTCTTCGTCTGCTTGGCTGCGACCGGGCGCAGGGTTTCCTGATTTCT
AAGGCAGTGTCTGCTCGTGAGTTCGAGCGGCAGTTAAGGGAGGACGGCCCCAGCCTCCTTGTT
TAA

Protein sequence: (SEQ ID NO: 140)

MSWKSRYRVLVVEDQPFQREYLLNLFRERGVQYLVGAGDGAEALRCLKQDRFDLILSDLMMPGM
DGIQMILQLPYLKHRPKLALMSSSSQRMMLSASRVAQSLGLSVIDLKPKPTLPKAIQGLEHL
ERCLRQKLEPETDETPHGRRTALLDALHNEQLVTFWQAKKSLHTGRIVGAELIRWSHPQHGLL
LPSCFMSDVEDATGLHEALLWRVLEQTLNAQESWRRAGYEIPVSVNLPPHLLDNQELPDRLYEY
VGARGACTSSLCFELTESSVTTLSSNYAGACRLRMKGFGLAQDDFGQGYSSFYNLVTTTPFTE
LKIDRSLVQGCVEDNGLNAAVISCIELGHRLLNDVVAEGVETCEELNLLRRLGCDRAQGFLIS
KAVSAREFERQLREDGPSLLV.

FIG. 17

RL038

DNA sequence: (SEQ ID NO: 17)

GTGAAGTCTGCTAGTGCCTTGGAGCACGACAACAACTTTTGCTCAAATGGACAACCTCTCGCAGAGCCTGAGCATCGGCTTG
 ATCTGTGTGGTGGTGTGCTGACCGTATTGCTGTTTACGATCTGTTACTGGTTCGCTGGGGAGATTGTTTTCAGGAGGAGGAGGACAAA
 GTCTCCTTCCACTTACCCGTATGATGGATGTTATACGGGAGCATGAGGTATTTCTTGGGCGCATCGCTCGAAAAAGCGACAAG
 ACCACCCAGAAGTACGACTATGACGTGGTGCCTTTGACGCGGCACTTGTGGCAAAGGAAAAACGGATTAGCGGTCTATGAGGGA
 CGGGAGTTTTCTTTGCTATGCCATTTTACTGGCTACCAAGCACGCGTTGAGCGCCGATTCTCGGGAGATCCGTTTTCGCTC
 GGTGTATTGCTCGCCAATTTCTACGGAAGCTTCTGGAGTGTTCGCGCTATCCCGCGCCAAGTTACTGATCTTTGATCTTTCC
 GGCAGCACCCGCTGGCAGTGCCGTCGATTCCCTCCACAGCGCAGCGTGACAGGTTGAGCGGAAGCTATCCGATGATAGTTCGAG
 CGCATTTCTGGCGCGCTTGCGCACCCGCGCGGTGGGGAGGACGCTCAGCGTGTCCATTGGATACGCGTGTATCGCTATCGCGAC
 TCGGCGCTGGAGATGTTGGGAGTGCCTCGGGTGGATCTGCCGGAACACTCTGGTGGCACGACGAGCCGAACCATCTGATCATC
 GCTGCGAGCCTGCTTGTATCTCAGGCGAATCAATGACTTCGAACAGTTGGTTGAGCGCCCGCATTCGATTCTGTACAGCCTGGTA
 TCGCCGAGTGGCGAGGTATTGCTCGGCGCGGCCCCCTGCGACCGGCTGAGGGATGGCCTGAACCTCACCCGACAGGGGGTTCGCC
 GTTCAACTGCTCAGCTCAGCGCTGAGAACGTTGGCTCGCGGTCTACCGAACCGACTACGGCAATTTCTTTCGCCACTCCCGGTGG
 CTGGTGGCAGGTCTGCTGCTGACCCCGCGCTGCTCCTGGCGGTTGGCTCGGGATGCGTTGGTACACCAGCAGCGTCTGTAAC
 CCGGTGTCATCGGGCGCACCGGCAACTGGTGGAGAGCGACACCTTCAGCCGGAAGCTGATACAGACCGCGCGGCTGGCTCTGGTG
 GTGCTGACCCAGGATGACAGCAACTGGTGAACCTGCAACCACTTGGCCGCCAGTGGCTGGGCGGGCCACCGAGATCCTTGGG
 CTGACTTCCAACCTGGAAGCTTTTCGATGCGCGTGGGAGGACGATACAGGAGACATCTGTATCCAGGTGCGTGGGCGCTATTGGCAG
 ACCGCTTTCGCGCGCACCCGCTATGCCGCGCACCGAGGCGGTACTGTGCGTATTCAACGACATCACGGTCCACTGCGAGGCGGAG
 ACCGCGCTGTCCAATGCGAAGCGAGCAGCGGATGCCGCCAGCCAGGCCAAGACCTGTTCTGGCCCCGATGAGCCATGAAATC
 CGTACTCCCTGTACGGTGTCTTGGCACCTTGGAGTTGCTCGACCTGACCACTGAAACGAGCGGCAACGCGCTACCTACGC
 ACCATCCAGAGTTCGTCTGCGACGCTCATGCAACTGATTAGCGATGTGCTGGATGTCTCGAAGATCGAAGCGGGGCGAGATGGCT
 CTGACCCTGGCCGCTTCAATCCGCTGGACCTAGTGCGGGAAGTGCTTGGCAACTTTGCCGCCAGCGCCATGGCCAAGGACCTG
 CAGTTCTATGCTTGCATCGACACCGAAGTGCCCGCGCAACTGATCGGTGACGTGACGCGGATTGCGCAGGTGCTCAATAACTTG
 GTGAATAACGCGCTGAAGTTCACCGATATCGGACGGGTGGTCTGCGCGTGAAGTTGCTCTCCCGCAATGATGGTTCGAGCCCTG
 TTGCAGTGGCAGGTGCGCCGACACCGGTATCGGTATCGCATCGCACACGAGCAGGAGCGCTTGTTCGAGGCGTTCTACCAGGTTTCG
 GGAGCGCACCATGCCGCGGACCGGGCTAGGACTGTGATCTGCTGGCATCTGGCGGAAATGATGGGCGGTACCTGCGAATG
 GTCAGCGAGACAGGGCTCGGCGAGCAGCTTCAGCCTGGTGTGCTGAGTTGCCCGAGGACGAAACAGTCCGGGCTGGCTGCGCGCGG
 GGGCTCTTGAATCCGCTTGGTCCATGTGCGCTCGCCGCTGCGGGAGCTAGCCGACAGCGTAGGGGCGTGGCTGAAAGCCTGG
 GGTGCAAGGTGAGCAGCGGCGAGGCGCGCTCCGAGCTGGAAGCTTGTGCTTCTGGAGCTGTGCCGATGGCGCGCGCGG
 CCTGCTTCTTCGCCCTGGCCAGGCCCCCGGGTGCAGCGCTCCATGGATGCGCCTTGCCAGCCGAGCTGCGTGAGGACGGCTGG
 CGTGTGCGCCTGCACAACCTGGCGGGAATCGGCCAGGCTGCGCGAGGCTCTGGGTGGCGATATCCCGAGCAAACGCGCGCA
 AATGCTTGCAGCGCTCGGGGAGACTCGACCTGGAAGTGTGCTGCTGCGGAGGACAACCCAGTCAACAGGCGCTGCTTCCGCGAG
 CAACTCGCGCAGCTGGGTTGTGCGGTGAGCCTTGCAGCGATGGGCGGAGGCGCTGCGAGCTGTTCCAGAGTGGTTCGCTTCCG
 CTCTGCTCAGCGACGTCAACATGCCGAACATGACCGGCTACGAACTGACCCAGGCGCTGCGCGAAGCAGGCGAGACGCTGCCG
 ATCATCGGCGTGACCGCCAAACGCTTGCAGAGAGGGCGAGCGCTGCGGGCAGTGGGAATGAACAGTTGGCTGGTGAAGCCG
 ATCACTCTGCATACCTTGATGAACCTGCTCAGTGAAGTGTGCTGCGCGAGGCTGCTGCTTCCCGCGAAGCGCGAGACCTCGGC
 CCGCCCGCGCAGCTCGACGACGCTCTCACCGCAGGTGCGGGAACGATGCGCGCGCTTTCTTTCGAGACCATGGGCAAGGAC
 CTGGAGGCGCGCGGCAAGCGATTGCGCGCAACGACCCGAAGGGGCTGCGAGGACCTGCATCGCATGGCCGGCTCCCTGGCG
 GTGATGCGTGCAGCAACGCTGGTGGTGTGTCAGGGCGCGGAGGAAGGCTGCTGGAGTTCGCGCTTGAATGTTCCGCGCTG
 GAGATTGGCGAGGTGCTCGTTCATATCGAGCAGGCGCTGGAGTTGTGAGAAAGACGGGCTGA

Protein sequence: (SEQ ID NO: 141)

VKSASALEHNDKLLKWTTLQSLSIGLICVVVLTVLLFSICYWSLGRFLQEEEDKVSFHFTRMMDVIREHEVFLGRIARKSDK
 TTQKYDYDVVPLQRHLLAKENGLAVYEGREFSFPMPFLATKHALSADSSGDPFSLGVLLANFYGSFWSVSAYPAPQLLIFDLS
 GSTRLAVPSIPSTAQRDLRSGSYPMIVERILARLRTRPVGEDAQRVHWIRADRYRDSALEMLGVARVDLPETLWWHDEPNHLII
 AASLLDLRRINDFEQLVERPAFDSYSLVSPDGEVLLGAAPATGLRDGLNLTROGVAVQLLSQPENGWLAHYRTDYGNFFRHSRW
 LVAGLLLTALLAGWLGMRYWTSVVNPVHRAHRLVESDTFSRTLITQAPVALVLTQDDQQLVTCNHAAQWLGGPTEILG
 LTSNWKLFDAHQVPGDICIQVGGRYLQTAFAATRYAGTEAVLVCVFNIDITVHCEAETALSNAKRAADAASQAKTLFLARMSHEI
 RTPLYGVLGTLLELLDLTLNERQRAYLRTIQSSSATLMQLISDVLVSKI EAGQMALTLAAFNPLDLVREVLGNFAASAMAKDL
 QFYACIDTEVPAQLIGDVTRIRQVLNVLNNAKLFPTDIGRVVLRVKLLSRNDGRALLQVQVADTGIGIAHEQQLRFEAFYQVS
 GAHHAGGTGLGLSICWHLAEMMGHLRMVSETGLSSFSVLLELPEDEQSGLACRPGLLKSA CVHVRSPVRELADSVGAWLKAW
 GCKVSSGEAAPSELETCLVLELLPMAAGPASSPWGPVRVRASMDAPCPPELREDGWVRVGLHNLGAGIQALAQALGGDIPEQTPA
 NACARSGRLDLEVLVAEDNPVNQALLREQLLEELGCRVSLAGDGRQALQLFDSGRFDLLSDVNMPNMTGYELTQALRERGETLP
 IIGVTANALREGERCRAVGMNSWLKPIITLHTLHELLSEFARAGVVLPAQARDLGPPAQLDDGLSPQVPERMRALFLETMGKD
 LEAARQAIRRNDPKGLQQLHRMAGSLAVMRARTLVVMCQGAEEGLLESRLCSAVEIGEVLVHIEQALEFVRKGT.

FIG. 18

RL039

DNA sequence: (SEQ ID NO: 18)

ATGCGTCCGGGGTCAATAGTTGGAATTAGAACACAAGAGAAGCCTATGAGTAAGCTCAAGATA
GTACTGGCCGATGACCATCCGATCGTGCGTATGGGCGTATGCGACATGCTCGAGCGCGACGGT
CGGTTTCGAGGTGGTGGGCGAGGCCTCCACGCCCAGCGAACTGGTCGAGGTGTGCCGGCAGAGC
GAGCCGCATATCGCCATTACCGACTACAGCATGCCCGGGGACGAGCGCTACGGCGATGGACTG
AAACTGATCGACTACCTGTTGCGCAACTTTCCTCGTACTAAGGTGCTCATCTTCACCATGGTC
GGCAACCGCCTGATCCTCGACAGCCTCTACGATCACGGGGTGTCCGGCGTGGTGCTGAAGAGC
GGCGAACTCGACGAGCTGCTCTTGGCGCTCGACGTGGTGAAGCAGAACCGCGTCTACCGGGGC
GCGAACATGCTCGACCCGACCAAGTGTCTGGCGAACCGCGACGAAGTGGAAAGCCGCTTCGCG
CGCTTGTCGATGAAGGAGTTCGAAGTACTCCGTCACCTTCGTTTCCGGCAGCAACGTCTGCGAT
ATCGCACGGCTGCTGAAACGTAGCGTCAAGACCGTAAGCACGCAGAAGGTCTCGGCGATGCGC
AAGCTGGAAGTGAACAGCGACCAGGCCTTGATGACCTTCTGCGTGCATGCCAACTTGTTCCAT
TGA

Protein sequence: (SEQ ID NO: 142)

MRPGSIVGIRTQEKPM SKL KIVLADDHPIVRMGVCDMLERDGRFEVVGEASTPSELVEVCRQS
EPHIAITDYSMPGDERYGDGLKLIDYLLRNFPRTKVLIFTMVG NRLLIDSLYDHGVSGVVLKS
GELDELLLALDVVKQNRVYRGANMLDPTSVLANRDEVESRFARLSMKEFEVLRHFVSGSNVCD
IARLLKRSVKTVSTQKVSAMRKLEVN SDQALMTFCVHANLFH.

FIG. 19

RL040

DNA sequence: (SEQ ID NO: 19)

GTGTCCAGTAAGATCCTGCTGCAAGGGGCACTGCTCGGCCTAGCAATGCTGGCCGTGCTGGAC
GCCCAGCCGGAGTCACCGCCGAGCGCACTCGGGCAATAATCGCCGAGGGGCACCGCGAGACG
TCGCTGCTGCTGGTCAACCAGAATGCCTATCCGGTCATAGTGCAGACCTGGATCGACGATGGC
GCCCCGAACTCGACACCGCAGTCTGCCCGCGCGCCGATCATGCCGCTACCGCCGGTGTTCGCG
CTCGAACCCGGACAGCAACGCAGCCTGCGCCTGCTGCGGACCGGCCAGGCGCTGCCAGGGGAC
CGCGAATCGCTGTACTGGTTGAACCTCTACGAAATCCCGCCGCAAGCCACCGGGCTGCTGGCC
GAAGGACAGTCACGGCTGACCGTTACACTGCGCACCCAGATGAAAGTCATCTACCGCCCTCGC
CCTCTGGCCAGAGGTGCGGAAGAAGCGCCACACCAGCTCAGGTTTCGAGCGGCGGGGCGAAACA
CTACAGATGGAGAACCCTACTCCCTATTTTCATCAGCCTCGCCGGCGCCGAGCTTGGCGGCCAC
ACCCGCCTGGCGGCGGCCGAACCTGTTGCCCCCTTCTCCAGGCGCGTCTTGGCGCTCCGCCAG
GCGCTGCCCGGCGGCCAGGCCGAGGTGCGCTTCAGCTGGATCGATGACGGCGGCAATCTCCAG
CAGGGACGGAGCCTGCTTCACTGA

Protein sequence: (SEQ ID NO: 143)

VSSKILLQGALLGLAMLAVALDARAGVTAERTRAI IAEGHRETSLLLNVQNAYPVI VQTWIDDG
APNSTPQSARAPIMPLPPVFRLEPGQQRSLRLRLRTGQALPGDRESLYWLNLYEIPQATGLLA
EGQSRLTVTLRTQMKVIYRPRPLARGAEEAPHQLRFERRGETLQMENPTPYFISLAGAELGGH
TRLAAAEELLPPFSRRVLALRQALPGGQAEVRFSWIDDGGLNQQGRSLH

FIG. 20

RL041

DNA sequence: (SEQ ID NO: 20)

ATGAAAACATCCCTGCGCGTCCTGCCTCTGCTCCTCGCGCTGCTCGCCTCGTCTAGCTGGGCG
ACCTGCTACAAGGTCACGGCGGTAGGCAACGCCACGACTACCTCCAACACCCAGATACGTCCC
GGTGAAGGCTCTGCCGGCACCTGGGCCGGAGCCTGCGATACCTGCAACGGTTCCCTCGGTCTA
CCGAGCGTGATCAACGTCAGCGACGCCAGCTTCCAGCCCCAGGTTAGCTTGATCGCCAGCTCG
GTGGCGCCGCTCAGCCAATACGGCGACAGCGCCGGCTACGACCCAGAGCGCGTGTCTTCCGC
TGTGCTCCAGAGGACGATGTGTACGAGATGTTCTCCACCAATGCCGACGATCTCTACAGCGGC
TGGTACCTAGGAGGCGACAGTGCGGGCAACTCGATTGGCCTGCAGTCCGCCTATCGCACCGCC
TGGCCCAACGTGCTGCTGCGCCTAACCCACGTGGAAACCGGGCAGTATTTACCGATGTCTGG
CGCGAGCGTCTGCTCGGCGGGCTCGATATCGACTCGCGAGGCTTTCAACTGGTCAAGGCGAAG
AACCTCAGCGCGGTACGCGCCGAACGTGTTCCGCGCTCCGCTGGAGTTCATCCGCTACTACTCG
CCGACTACCGCCTCGCGGTTGTACGCCTACACCCAGCCCCGCTGGCTACATCGCCATCAAGGGT
CCCGGCCTGGCCTACCCCAACGTGCGCGCCAGCCATAACGCCAACTACCTCGGCTGGCACTAC
AACTGGCCCGGGCGCCATCGGCCTGTACAACGACGTGACGCTCAAGCGCTATCCACCTGTTCC
GTAACCAACGTACGCCCCACGTTGTGTTCCCGTCGATTTCCCTCAGTGAGATTAATGCCGGC
GCGAACCGTGAGATGCCCTTCGAGGTGGCCTTCAAGTGCCAAACGGGAGTGATCAACAGCACC
GCCTCCAGCGGTACTGCACTGGGTATCAGGGCTTCAGCCGGGGCGCAGGCCGCGTCCGCTGCA
CTGGGCCTGAGGAACGCCAATGGCGGGCTCTCCTACCTAGTTTCCGACCGCTACGGCCAGCCT
GGTATGGCCCAAGGCGTGGGTATCCGCTTGCTGCGCGACGGCAGTGCGATGAACCTGCTGGTA
AGCGAGGATTCCGCGATGGGCAGCAATGCCGAAACACGGGGCTGGTATCCAGTGATCGGCAAC
GCCTCGAACAAGACTGGCGAAGCGGGAGGCATCAGCCAGTACAGCGAGACCTTCCGTGCGCGC
CTGGAAAACTCAACGTTGGCAGCATGCCCAGCGTTACCCCGGGACGGGTGGAGGCCAGCGCG
CAGGTAGTGATTCTGTGTCAGTAA

Protein sequence: (SEQ ID NO: 144)

MKTSLRVLP LLLALLASSSWATCYKVTAVGNATTTSTNTQIRPGECSAGTWAGACDTCNGLGL
PSVINVS DASFQPDGSLIASSVAPLSQYGDSDAGYDPERVFFRCAPEDDVYEMFSTNADDLYSG
WYLGDSAGNSIGLQSA YRTAWPNVLLRLTHVETGQYFTDVWRERLLGGLDIDSRGFQLVKAK
NLSAVRAELFRAPLEFIRYYSP TTA SRLYAYTQPAYIIA IKG PGLAYPNVGASHNANYLGWHY
NWPGAIGLYNDVTLKRYPTCSVTNVTPHVVFPSISLSEINAGANREMPFEVAFKCQTGVINST
ASSGTALGIRASAGAQAASAALGLRNANGGLSYLVSDRYGQPGMAQGVGIRLLRDGSAMNLLV
SEDSAMGSNAETR GWYPVIGNASNKTGEAGGISQYSETFRARLEKLTVGSMPSVTPGRVEASA
QVVIRVQ

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND
PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September 12, 2003 Serial No.: Not Yet Assigned

Page 36 of 118

Customer No.: 21559

FIG. 21

RL042

DNA sequence: (SEQ ID NO: 21)

ATGTTCTGCCACGTTGAGGCACGGCGCACCGGCAAAC TGCCGCTGGCTCTGGGCGGCCTAGCGCTGGCCTTCGCGG
GCCTGGCCAATGGAGAGGCGCAGTATCGTTTTCGACGACAGCCTGTTGATGGGCTCGGGCCTCGCCGGCGGGACCC
AGAAGCCTTCAACCGGGCCAACAGGTGGACCCCGGAACCTACCATGTCGATGTCTATCTCAACGGCAGCTACGCC
AGTCGCACCAGGATCGAGTTCCGCCCCCGGGCCGGCGGCGTCAAACCCCTGCTTCGGCGAACGCTTCTTGGCGCCGA
CGCTGGGCGTCCGCCCCGCTCTGAGGCGGGCGTGCAAGCGCCTGGAGATTGCCTGGGGCTGGAAGAAGCCTGCC
AGGCTCGACCTTCAATCTCGACACCGCCCTTCTGCGCCTCGATCTCTCGGTGCCCCAGGCCCTGCTGGATATCAAG
CCACGCGGCTACGTGGGTCCCGACGAGTGGGACGCTGGCAGTAGCATGGGCTTCGTCAACTACGACGCCAGCTTCT
ATCGCTCGAGCTTCGACGGAGTAGGCGGCAACGGCGACTCGGACTATGGCTACCTGGGGCTGAGCGGGGGCATCAA
TTTCGGCCTGTGGCGCCTGCGCCACAGTCCAACCTACAGCTACTCCAGCTATGCGGGAACACCCGACGCGACTGG
AACAGCATCCGCACCTATGCCCAGCGCGCGGTGCCAGGCCCTGCGCAGCGAACTGACCCTGGGCGAGAGCTTCACCG
AGGGCAATCTGTTCCGCGAGCCTGGGTTATCGCGGCGTGCGCCTGGCCAGCGACGACCGCATGCTGGCAGACTCGCA
ACGCCGCTATGCTCCACAGGTACGCGGTACAGCAACAGCAACGCACGGGTGGTCACTAGCCAGAACGGCAAGAAG
GTCCACGAATCCGCCGTCGCTCCCGGTCCCTTCGTATCAACGACCTCTATGGCACCGCCTACGACGGCGATCTGG
ATGTCCAAGTGATTGAGGCGACGGCAGCGTCTCGCGCTTTTCCGTGCCCTTTTCCGCGGTTCCCGAATCCATGCG
CCCGGGCATCTCGCGCTACAGCGCCACCCTCGGCCAAGCGCGCCAGTATGGCGACGGCAACGACCTGTTCCGGCGAC
TTCACCTATCAGCGCGCCTGACCAACTCGCTAACCGCCAACCTCGGCTCGCGCCTGGCCGAGGACTATCTGGCGC
TGCTCGGCGGAGGCGTGCTCGCCACGCCCTACGGAGCCTTCGGCTTCAACAGCATCTTTTCCCATGCCACGGTGGA
GAACGGCCAGCGCAAGCAGGGCTGGCGTGTCGGTCTGAACCTACAGCCGACCTTCAGCCGACCCAGACCACTCTC
ACCCTGGCTGGCTACCGCTATTCCACCGAGGGCTATCGGACCTCGGCGACGCGCTTTCGGCGCGCCACGCCGATG
AGCACAACTCCTGGAACCTCAGCAGCTACAAGCAACGCAACCACTTCAACCTGCTGGTCAACAGGGCCTGGG
GGGCTACGGCAACCTGTATCTGTCCGAGCCACCAGCGACTACTACGACGGCAAGAGCCGCGACACCCAGTTGCAG
TTCGGCTACAGCAACACTTGGCGCCAGCTCAGCTACAACCTCGCCTATTTCGCGCCAGCAGACCACCTGGTACCGCG
ATCTGAACGACGACTACGACCCGCTCACTGCCGCGCGAATACAACCTGCGGCACGGCAGCGAACGTAGCAACACCTT
AACCCTGACACTTTCATGCCGCTGGGGTCTCCAGCCAGGCCCGCAATCTCAGCGCGATGGCCTCCCGGCGTTC
GGCGACAGCCGCGGCGCAGCTACGACAGGGGCTCAACGGCACCTCGACGAAGACCGCAGCCTGAGCTACGCGCA
TTGCCGCGCGGCGCGACGACGACAACACGGCAGCGATTTCAACGGCAGTCTGCAGAAACAGACCTCGGTGGCGCA
GCTGAACGCCGCTATGCCGAGAACAGCAGCTACCGGCGCTCAACACCGGCTGCGCGGCGCGCGCTGCTGCAT
CGCGGCGGCTGACCTCGGCCCCCTACGTCGGCGACACTTTCGCCCTGGTTCGAGGCCAAGGGCGCCAGCGGAGCTG
GCGTACGCGGTGGTTCAGGGCGCGCGCGTCAACGGCAATGGCTACGCCGTGGTGCCATCACTCTCGCCCTACCGCTA
CAACCCGCTCAGCTCGATCCGCGAGGGCATGGGCGAAGAGGCCGAGCTGCTGGAGACCGCGCAAGATCGCGCCA
TACCCGCGGCGCGCGGCTGCATGTGAAGTTCCGACACTGACCGGTACCCATTGCTAATCAGGCCCACTCGCCG
ACGGCAGCGCGCTACCGCTAGGGGCCAATGTGCTCGACAGCCAGGGTGTGAACATCGGCATGGTTCGGTCAAGCGG
CCAGGTCTATGCCCGCGCGGAGGGCGACAAGGGCCGCTGCGCGTGCAATGGAGCGAACGCCAGGGGACGCTGT
CTGCTGGATTACGACCTCGACACTGGCCCTCGCCAGGCTATCGAACCCGACAGGCGGTGATCCGCTGCAGGGCA
CCTGCACGCCGCTCTCGGAGGCACCATGA

Protein sequence: (SEQ ID NO: 145)

MFCHVEARRTGKLPALGGLALAFAGLANGEAQYRFDDSLMGSGLAGGTLEFRFNANQVDPGTYHVDVYLNGSYA
SRTRIEFRPRAGGVKPCFGERFLRRTLGVPRASEAGVQAPGDCGLGLEERLPGSTFNLDTALLRLDLSVPQALLDIK
PRGYVGPDEWDAGSSMGFVNYDASFYRSSFDFGVGGNGDSYGYLGLSGGINFGLWRLRHQSNSYSYAGNTRSDW
NSIRTYAQRAVPLRSELTLGESFTEGNLFGSLGYRGVRLASDDRLADSQRRYAPQVRGTANSNARVVISQNGKK
VHESAVAPGPFVINDLYGTAYDGDLDVQVIEADQSVSRFVSPFSAVPESMRPGISRYSATLQARQYGDGNDLFGD
FTYQRLTNLSLTANLGSRLAEDYLALLGGGVLATPYGAFGNISIFSHATVENGRQKQWRVGLNYSRTFQPTQTTL
TLAGYRYSTEGYRDLGDALSARHADEHNSWNSSSYKQRNQFTLLVNQGLGGYGNLYLSGATSDDYDGSRTDQLQ
FGYSNTWRQLSYNLAYSRRQTTWYRDLNDDYDPSLPQYNLRHGSESNLTLLTSMPLGSSSQAPNLSAMASRRS
GDSRGSSYQTLNGLTDEDRSLSYAIAAGRSDSNHGSDFNGSLQKQTSVATLNAGYAENSSYRQLNTGLRGA AVLH
RGGLTLGPYVGDTFALVEAKGASGAGVRGGQGARVNGNGYAVVPSLSPYRYNPVSLDPQGMGEEAELETERKIAP
YAGAAVHVKFRTLTGHPLLIQAQLADGSALPLGANVLDSQGVNIGMVQGGQVYARAEGDKRLRVQWSERPGDAC
LLDYDLDTGPRQAIEPGQAVIRLQGTCTPVSEAP

FIG. 22

RL043

DNA sequence: (SEQ ID NO: 22)

ATGAATACTTTTCCACTGCCTCCGCTCCGTGCGGCTACGCTGGCGCTCGCCCTGCTGATACCC
GCCATCCCGGCTCAAAGCAGCGTGGTGATCATCGGTACTCGCGTGATTTATCCCGGCGACGCC
CGGGAAAAGACCGTGCAGATGATCAATCAGGACGCATTCCCCAACGTGATCCAAGCCTGGATC
GACAACGACGACCCCTCCTCCACCCCGGAGACTGCAAACGCGCCCTTTCTGGTCAGCCAGCG
GTGACGCGCATAGCCCCCGGCAGCGGCCAGACCCTGCGCCTCCTGTATACCGGGCTCCCGCTG
CCCGAGGATCGCGAATCGTTGTTCCATCTCAATGTGCTGCAGATCCCGCCCCGCGACCTGGCC
AAGGCCGAGCGCAACCAGATGCTGCTGATGCAGCGCAGTCGACTGAAGCTGTTCTATCGCCCC
GCCGCGCTGCTTGGCGGCTCGGAGCAGCTAGTCGAGCAGTTGCACTTCAGCCTGGTGCAGGCG
AGCGGCAACTGGCGTGTGCGGGTGGACAACCCAGCGGCTACTACGCCTCCTTCGCCGGCGCG
ATGCTGAGCATCGGCGAACGTCGCTGGCGGCTGCTGTGAGCATGGTCCCGCCCCAAAGGCCAG
GCCGAGTGGGCGGCGGAACGCCCTTCGCCGCTCGCCCCAGGACCGGTCCAGTTGAACGCCCTC
TTGATCAATGACTACGGCGCGCGAATGGAGGTCCAGCATGTTCTGCCACGTTGA

Protein sequence: (SEQ ID NO: 146)

MNTFPLPPLRAATLALALLIPAIPAQSSVVIIGTRVIYPGDAREKTVQMINQDAFPNVIQAWI
DNDDPSSTPETANAPFLVSPAVTRIPGSGQTLRLLYTGLPLPEDRESLFHLNVLQIPPRDLA
KAERNQMLLMQRSRLKLFYRPAALLGGSEQLVEQLHFSLVQASGNWRVRVDNPSGYASFAGA
MLSIGERRWRLLSMVPKQAEWAAERPSPLAPGPVQLNALLINDYGARMEVQHVLP.

FIG. 23

RL044

DNA sequence: (SEQ ID NO: 23)

ATGAAACCTCAAAGTACTGCCCTGACTATCGCCGCATTTCTCGCATTGCCGGGTATCGCGGCG
GCTGCCAATACCATCACCTTCCACGGAGAAGTGACCGACCAGACCTGTTCCGCCGTCGTCGAC
GGACGAACCGACCCGACCGTGATACTCGACACCGTACCGGTAAGCGCTCTTGACGGCGCAGTC
GGCAAACCCGCCGGGGAAACCAGCTTCACCCTGCAACTGACCGGTTGCGCCGCTCCGGCGGCC
GATGCCGAGGAGCACTTCAGCGTGATGTTCCAGGCGGTCAATCCGACCAGCGCCGGCAATCTG
ACCAATACCGCGTCCGCCGGCGCCACCGGCGTAGCGCTGCAGCTACTGACGGCACCGGGCGGC
AGCGAGGTCAATCTGGCCGGCGGGTCGGCCGTGGCTGCCGGTGACATCGTGCTCGCAGGAGGC
GAGACCAGCACCAGCTACGACTATGCCGTCCGCTACATCTCCGAAGCGACCACCGTCACTCCG
GGACCGGTGCTCGGCTCGGTGACCTACACCCTGCGTTACGAGTAA

Protein sequence: (SEQ ID NO: 147)

MKPQSTALTIAAFLALPGIAAAANTITFHGEVTDQTCSAVVDGRTDPTVILDTVPVSALDGAV
GKPAGETSFTLQLTGCAAPAADAEEHFSVMFQAVNPTSAGNLTNTASAGATGVALQLLTAPGG
SEVNLAGGSAVAAGDIVLAGGETSTSVDYAVRYISEATTVTPGPVLSVTTYTLRYE

FIG. 24

RL045

DNA sequence: (SEQ ID NO: 24)

AGTCCGCACGGTAGTGACGACTGGAAGCGCTTCTGTGCTGCCAACACCTGGAGCCCAGCATG
AGCCGGCGCGGCAATTGTTGGGATATGCCGTGGCGGAATCCTTCTTCAGTAGTTTGAAGAAAG
AGCGTATCCGCAAACGCATCTACAAAACCCGAGACATGGCCCGGGCGGATGTTTTTGA CTACA
TCGAGGTCTTCTACACCCGAACCCGGCGGCACAGTCATCTGGGTGGCGTCAGTCCCGAGGCCT
TTGAAAGCGCCTCG

FIG. 25

RL046

DNA sequence: (SEQ ID NO: 25)

ATGGCTGAAGTCACTCAACGAGCAGAGCAGCAACAAGAGAGCCAGAAGACCCTTCTCGGCACC
ATCATCAGTACGCCCTTCCAATTTCTCGGCGTGATGTTCTGGGTCGCTGATCGGCGCAATCATC
GTGGAGTGGGTTTGCCTGTATTTCTTCTGGCCTGACGCGGGCTGGAAGCATGCCCAGGCCATG
TTTGAGTACGAACTCAGTTGGCTGTTCGAGGGGCTGCTACACAGCGTCGTCGTGCAGGAGCCA
GGTCGAACCGCCACCTGGCTGGCCCAGTTGGCCTATGACTGGTTGTTTCGTGAAGACCGGGATG
GTCGACTGGATGACCAACATGACTACCATCGCGCAGGCCGGGCCACGGAGCCCGCTGGACGTT
CGCTATCTCACCGCCAGGGTGTCTCCACGCTGCAGAACTACGGCCTGGCCGCGCTGTACACG
GTGCTGACATTTCGTTCGTGCGCCTGGTGATCCTGGTCATGACGATCCCGTTATTCGTGATGGCC
GCGTTCACCGGCCTGGTGGACGGCCTGGTGCGCCGGGACCTGCGCAAGTTCGGCGCCGGCCGG
GAGTCCAGCTACCTCTACCACAAGGCGCGCGGCAGCATCATTCGCTAGCGGTTCGTCCCTTGG
ACGCTCTACCTGGCAATTCCTCATCAACATCAATCCCCTGCTCATCCTGTTGCCCTGCGCCGCA
CTGCTCGGCGTAGCGGTGTGCATCACAGCATCCACCTTCAAAAAGTACCTATAG

Protein sequence: (SEQ ID NO: 148)

MAEVTQRAEQQESQKTLTGTTISTPFQFLGVMFGSLIGAIIVEWVCLYFFWPDAGWKHAQAM
FEYELSWLSQGLLHSVVVQEPGRATATWLAQLAYDWLFVKTGMVDWMTNMTTIAQAGPRSPLDV
RYLTAQGVSTLQNYGLAALYTVLTFVVRVLVILVMTIPLFVMAAFTGLVDGLVRRDLRKFGAGR
ESSYLYHKARSGIIPLAVVPWTLYLAIPININPLILLPCAALLGVAVCITASTFKKYL.

42/118
FIG. 26

RL047

DNA sequence: (SEQ ID NO: 26)

ATGGCTGGCCAGTACCCGTTGGAAGCGCTCTTGCGGCCTGCCGTGGAGCTCTACACCACCACCGTGTGC
TTCACCGCAGCCGCGCTCTGCATCGTCGCGCCGTGGACGTTCTCCCTCACTCCGCTGTTTCGGCATCGTG
GCCGCGCTGTGCTTCGCCTGGCTGGGTATCGTGCGGCTGAAGCAGGCCGCGGTGGTGCTCCGCTACCGG
CGGAACATTCGCCGACTGCCGAAGTACACGATGACCAGCGCCGAGATGCCGGTCAGCAACGAACACCTG
TTCATCGGTAAAGGATTTTCGCTGGACGACGAGCAGATACGACGCGCTGGCAGATACCTACCTGCCCCAG
TTCGCCTCTTACGTCGAGCCCTCGCCCCCTCTACGAGCGCGCGCGCCGGTTGGAGAAGCAGCTCGAGTTC
GCCCCCTTCCCCCTGAAGCTGGTCGCCAAAGCCACTGCCTGGGACGTGGCCTGGAACCCCGCACGGCCG
CTGCCGCCCCGTGGGCGGTTTGCCTCGGCTCCATGGCATCGAGCCGCGCGAACAGGACGTAGGCCTGCAA
CTGGGCGAGCGCGTCGGCCACACACTGGTACTCGGCACCACGCGGGTGGGTAAGACGCGCCTCGCGGAG
CTGTTTCATCACCCAGGATATTCGCCGCACTCACTGCCGGGTACGACGCCGCGGGTGAAGATGGGCGG
CGGACCCAGACGGTTACACACGGCTATCGCGCCGCGCGCGCAGAGGAGCAGCCGGATACGAGGTGGTG
ATCGTCTTCGACCCGAAGGCGACGCTGACCTGCTGAAGCGTATGTACGTGGAATGCGAACGTCGCCG
CGCCTGGACGAGTTCTACGTGTTCCACCTCGGTCATCTGACCTGTTCGGCACGCTACAACGCGCTCGGC
CGGTTTCGGTCGGATCTCCGAGGTCGCCACCCGCGTCGCCGCGCCAGCTCTCCGGCGAGGGCAACAGCGCG
GCGTTCCGCGAGTTTCGCTGGCGGTTTCGTCAACATCATCGCCGCGCGCTGCACGCGCTGGGTATCCGG
CCTGACTACCAGCAGATCCTCCGGCACGTCGTGAACATCGATGCGTTGTTTCGTGCAATATGCGCAGAAA
TACATCAGCGAGCACGATCCAGGGCCTGGGACACCATCATCCAGATCGAGGGCAAGCTCAACGACAAG
AACATCCCGTTCAACATGAAAGGACGGCCCCCTGCGGGTCGTAGCCATCGACCAGTACCTGACACAGAAA
CGCATCGCCGACCCGGTCATGGAAGGCTTGAAGAGCGCCGTGCGCTACGACAAGACCTACTTCGACAAG
ATCGTGGCCTCGCTGCTGCCGCTACTGGAGAACTCACTACCGGGCGGATCTCGGAGCTTCTTTTCGCCC
AACTACGCGGACCTCAACGATCCGCGGCCGATCTTCGACTGGATGCAGGTTCATCCGAAACGCGCCGTG
GTCTACGTCGGCCTCGACGCACTATCGGATACCGAGGTCGCCGCCGCGGTGGGCAACTCCATGTTTACG
GACCTGGTCTCGGTAGCGGGTCACATCTACAAGCATGGTGTTCGATGACGGCTGCCCGGCTCGCTCGCC
AGCGGCAAGGTCGCGATCAACCTGCATGCCGACGAGTTCAACGAGCTGATTGGCGACGAGTTTCATCCCC
ATGGTCAACAAAGCGGGCGGCGCGCGCGTGCAGGTGACGGCCTACACCCAGACCATGAGCGACATCGAG
GCCAAGATCGGCTCCCGCGCGAAGGCCGGTCAGATCATCGGCAACTTCAACAACCTGTTTCATGCTGCGG
GTGCGCGAGACCGCCACGGCCGAACCTCCTTACCAATCAGCTCCCCAAGGTCCAGATCTACACCAGCACG
CCGGCGAGCGGCGCCAACGACGCGATCAACAACAACAAGAAGGTAGCCTTCACCTCCAGCTCGCACGAC
CAGGTGCAGATGACCAGCGTGCCGATGCTCGAGCCGCGCCACATCATTGGTCTGCCCAAAGGACAAGCG
TTCGCGCTACTCGAGGGCGGCAATCTCTGGAAGATCCGAATGCCGCTGCCGCGGTCGCCCCCGACGAG
GTGATGCCGAAAAGCCTGCAGGAGCTGGCTGCCGGTATGCGCAAGGGCCAGGCCGCAACAGCGAGTGG
TGGGAGGCGCCGGGATACTCCGCCCTGCAGGATGGTCTGCCCCAGGACCTGGTCGACGATTTCCGTCAC
CTCGGCACCGGTGAGGATGCCGCTGA

Protein sequence: (SEQ ID NO: 149)

MAGQYPLEALLRPAVELYTTTVCFATAALCIVAPWTFSLTPLFGIVAALCFAWLGIVRLKQAGVVLRYR
RNIRRLPKYTMTSAEMPVSNEHLFIGKGRWTQKHTQRLADTYLPQFASYVEPSPLYERARRLEKQLEF
APFPLKLVAKATAWDVAWNPAPPLPPVGGLPRLHGLEPREQDVGLQLGERVGH TLVLGTTRVGKTRLAE
LFITQDIRRTHCRVRRRRVKMGRRTQTVHHGYRRRRRAEEQPDYEVVIVFDPKGDADLLKRMVVEECERAG
RLDEFYVFHLGHPDLSARYNAVGRFGRISEVATRVAGQLSGEGNSAAFREFAWRFVNI IARALHALGIR
PDYQQILRHVVNIDALFVEYAQKYISEHDPAWDTIIQIEGKLNKNI PFNMKGRPLRVVAIDQYLTQK
RIADPVMGLKSAVRYDKTYFDKIVASLLPLEKLTGRISELLSPNYADLNDPRPIFDWMQVIRKRAV
VYVGLDALSDTEVA AVGNMSFSDLVSVAGHIYKHGVDDGLPGSLASGKVRINLHADEFNELIGDEFIP
MVNKAGGAGVQVTAYTQTMSDIEAKIGSRAKAGQIIGNFNLFMLRVRETATAELLTNQLPKVQIYTST
PASGANDAINNNKVAFTSSSHDQVQMTSVPMLPAHIIGLPKGQAFALLEGGNLWKIRMP LPAVAPDE
VMPKSLQELAAGMRKGQAANSEWWEAPGYSALQDGLPQDLVDDFRHLGTGEDAA

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND
PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September 12, 2003 Serial No.: Not Yet Assigned

Page 42 of 118

Customer No.: 21559

FIG. 27

RL048

DNA sequence: (SEQ ID NO: 27)

ATGACTACTCATCTGATCACCCCTAGTCATCAAGCAGCCGAGCGACGCTCAGGCACGCCAACTC
ATGTACCAGGAGTTGCTCGGACTGATCTCACGCTACGGCGGTGAGGTGACGTCCAAGGCCTTG
GAGGACGAGTCGACCCTCTGCGAGCTGCTGGTGCAGATGCTGCCTGATCATGAGGTAGAGCAA
GCCAGGAAACAGGTGCTCGAACTTCATGCCAAGGGCCGCCTGCAGGCGCCGGCAAGCCTGAAG
GTGTAA

Protein sequence: (SEQ ID NO: 150)

MTTHLITLVIKQPSDAQARQLMYQELLGLISRYGGEVTSKALEDESTLCELLVQMLPDHEVEQ
ARKQVLELHAKGRLQAPASLKV

FIG. 28

RL049

DNA sequence: (SEQ ID NO: 28)

ATGAAGAAGTTCCTTGCCACGCTGGCATT T TGCACGGCGTTCGCGACTCAAGCCTGGGCCGCC
GGGCTGATCGTTGTCTGAAGACCTCGGCGGCGCCTCGGCGCTCCCCTACTACCAGGGCCTGGAT
CCGCAGCCATCCGCTTCCGCACCAGGACCTGGCGACCTGGGCGTCCGTGGCTCAGGTGCGTTT
CCAGTTCGCTCCGCCCCCCTATCGCCAGGACGGGTCCAGGGGCGCGCCATCAACGCTCCAGGC
CTGCAACTGCTGTTCTTGGTTCGGCGACGACACGCTGTCTCGAACCTGGCTGAAAGAGCGAGGC
GACGAGCTTCGAGACCTCCAAGCCGTGGGCCTGGCAGTGAACGTGGCCAGCGAAGCGCGCCTG
ACGGAATCCGGGCCTGGGGGAAAGGACTTCAGATATTGCCGGCGCCGGCGGACGACCTGGTC
GACCGGCTAGGGCTGCAGCATTACCCCGCCCTCATCACATCCACCGCCATCCAGCAGTAG

Protein sequence: (SEQ ID NO: 151)

MKKFLATLAFCTAFATQAWAAGLIVVEDLGGASALPYYQGLDPQPSASAPGPGDLGVRGSGAF
PVR SARLSPGRVQGRAINAPGLQLLFLVGDDTL SRTWLKERGDEL RDLQAVGLAVNVASEARL
TEIRAWGKGLQILPAPADDLVDR LGLQHYPALITSTAIQQ.

FIG. 29

RL050

DNA sequence: (SEQ ID NO: 29)

ATGGCAACGTCTGTAGTTCGAGCCCTCCAGTTGGCCACCCTGCTGGTCCTGGTCAACATCGCT
CAGGCCGCCGTGGATCCACCGCCGGCGTACAAGCAAATCGCCCTGCCCAAAGGGGTTCCGGCC
GAGGTGCTCTACTCGGTCGCGCTGACCGAGAGCAAGGTCCTGCTGCGCGGC GAATACGTTCCC
TGGCCCTGGACATTGAACGTCGCCGGGAAATCTTACTACTACGCGACCCGCACCGCCGCCTGC
ACAGCGCTACTCGCGGCGATCAACCTCTACGGGGCCAAGAGCGTCGATTCCGGCCTCGGCCAG
GTCAACATCGGCTGGAACGGACATCGTTTCTCCAGCCCCTGCGAGTCCCTGGATCCGTACAAG
AACCTGGACGCCACCTCCGACATCCTGATCGAGCAGCGGGACGCCCTGTATGCATCCGCCCCG
GGAAGACCGGTGGACTGGATCCAAGTTGCCGGCCGCTACCACCGCCCCGCCGGCGCGCCT
GCCGCCAAATACCGTAGGACGGTTTCCCGCCACCTTAGCCAAGTTCTCGGCGTCAACCTACTG
GTGACCAATCCATGA

Protein sequence: (SEQ ID NO: 152)

MATSVVRALQLATLLVLVNIAQA AVDPPPAYKQIALPKGVP AEVLYSVALTESKVLLRGEYVP
WPWTLNVAGKSYYYATRTAACTALLAAINLYGAKSVDSGLGQVNIGWNGHRPSSPCESLDPYK
NLDATSDILIEQRDALYASAPGRPVDWIQVAGRYHRPAGGAPAAKYRRTVSRHLSQVLGVNLL
VTNP

46/118
FIG. 30A

RL051

DNA sequence: (SEQ ID NO: 30)

ATGATCAGAACCGTATCGCTCCTGTCCGGCCTGATGCTGCTGCTGAGCTATCCCGCAGCCGGCCAGGAGCGGCGGCAAGCCG
AGAGGCCAGCAGCCAACTGTCCGGTAGCCAACTCGGCACGCTGAAACAACAGACATCTCAGAGCGACCTGGCCAGGAGTGGG
GACTGAACCAACAGGAATGGACCCGCTACCAGACGCTCATGCAAGGCCCGGGCGCTTACTCGCCTGGTATTGATCCGCTG
ACCGCGCTGGGCATCGAGGCGCGATCGGCAGAGGAACGGCGCGGTATGCCGATCTACAGGTCCAGGCCGAACGGCGCCGGT
CGAGAAGGAACTCGCTACCAGCGCGCATACGACGAAGCCTTCGCCCGCGCTATCCAGGCGAGGGGGTGATCCGCTCACCG
AAAGCAGCACAGCCAACCCGTCGGGCACGCCGAACATGAGCCAGCGTTGCAGAGCAGCGGGCGCCTGGCCCTGTTCTGTCAG
GACAACTGCACCGCCTGCATCCAGCGGGTCCGCGACCTGCAACATGCAGAAAAGGAGTTTCGACCTCTACTTCGTCCGTAGCCA
GAACGACGCAGAGCGAGTGCAGCGCTGGGCAATCCTCGCCGCGATCGACCCGAAGAAGTTTCGCAGCAAGCAGATCACGCTCA
ATCATGACGAGGGCCGCTGGATGGCCCTAGGACTGGGCGGAGCCCTTCCCGCCCTGGTACAGGAGGTGAACGGCCGATGGCAA
CGTCTGTAG

Protein sequence: (SEQ ID NO: 153)

MIRTVSLLSGLMLLLSYPAAGQEAASREASSQLSGSQLGLTKQOTSQSDLAQEWGLNQEWTRYQTLMOGPRGAYSPGIDPL
TALGIEARSAEERRRYADLQVQAERRRVEKELAYQRAYDEAFARAYPPEGVIRLTESSTANPSGTPNMSPALQSSGRLALFVQ
DNCTACIQVRDLQHAKEFDLYFVGSQND AERVRRWAILAGIDPKKVRSKQITLNHDEGRWMALGLGGALPALVQEVNGRWQ
RL.

RL052

DNA sequence: (SEQ ID NO: 31)

ATGAAACGCCCATCCCTGCATCAATGATTCTTGGCCTCTGTTTGACGGCAATGGCCGGCCTGCTGAGCTACCAGCAGTACCA
ACTCGTTTCACTCCGATCAGGCGTGGACAGTGCCCGGAAAAGGCCTCGCTGGAGGCGATCCTGGCTCGCTTGAGTCGAGTCG
ACGAGCGCCTCGACGCCGTGGATGGACAGACCTGGTCAGCAACGAGGACTTCCGTTTCAAGGCGAGGCGCTGTCCAACCGA
ATTGACGCTGCGCAGGCGTTTCGCCAAGCAGGCCTCCGATGCCGTGAGAACCTGGCTCAGACCACCGCCTCGGCCGGCGACCT
CTTGGTGCTCAAGGCAACCGTGGAGACACTGGACGGTTCTGTCCGACGCTTCAAGAAAAGCAGGCCAAGGCGCGCCGCTGA
TCGTGCCAGCGCCAAAACGCCCATACCCGCCAAGCCGAAACCCAAACCGATGGAGCCCCCGCCCTTCTCGATCCTT
GGCGTGAGTATCGCGGGGAGAACGGTTTCTGTCCGTTGCACCTCCGGGATCCACCCAGCTCAGCCAGATCTACCTCATTCG
CCGGGGAGATGCCGTGCGCGGCACGACCTGGCGACTGACCGACCTTGACGATGGTACCGCGCACTTCGACGTCGCCGGCACCT
CGCGCAGCGTTTCGATCCAACCATAG

Protein sequence: (SEQ ID NO: 154)

MKRPSASMILGLCLTAMAGLLSYQQYQLVQLRSGVDSAAEKASLEAILARLSRVDERLDAVDGQHLVSNEDFRSGQQALS
IDAAQAFQASDAVENLAQTASAGDLLVLKATVETLDGVSRTLQEKQAKAPLIIVPAKRPPIPAKPKPKPKMEPPPPFSIL
GVEYRGGERFLSVAPPGSTQLSQIYLIRRGDAVAGTTWRLTDLDDGTAHFDVAGTSRSRVRIQP.

RL053

DNA sequence: (SEQ ID NO: 32)

ATGCCGCGCCGCTTGATCCTCTCGGTACCGGAGCGGGATATCCTATTTGCACTGCCGGTAAGCCGAATGACCTCACTGACTAC
TCCACCTCAACGAGTCCGCCCCATCGTCGATCCGCCAGCGACGCGGCGATGCCAATCACTTGGTTTTTTCGGTGCAGGTCAGC
CTGCTGTGCTATCCAGCGTTTACGCCCTGATGCGCGACGAAGAGCCCCCGAG

RL054

DNA sequence: (SEQ ID NO: 33)

ATGGCCGAAGCTATCAGAAAAGGATGCAATGATGACAAAACCTCTACTTTGATCTTCTGAACTCGCCTGCCGAGGCTCATTTCGTC
GATACAAAAGTCTTTATCTGTGCAGGCAATCTCCACAACCTGTCCCAATACTGGAGTTTCTTTCGAAACCGTATACGCCCTATG
CATCGTACATAAATGCATTAAGTATCGGTCAACGCATAGATCCTGCATTACCCAGAGCTTAACGAGTGCCATATCCAACCTG
GCAGGTGCGCCGATTGCAGTAAGCGACATTTACAAAAAATTATGAAACCACACTGAGAACACCTGTTGAGATGGGCGTTTCG
TCCTAATAGCATCACCTTTGAGGAGTATCAGGCCACCATAAATCAGCAAGCCATCAACATGGTTCAAGATATGCAGGATGGAG
ACAAAGGTGAGAAGGTGGAGGCCCTCCAGGCCAATATGCAGTTCTGTATGGACAGGAGATAAATACTGATTTTCATCGCTCGT
AATGAACTCGCTGCTGGGCAGAGAGCGAAAACCGTCGCAATAGTTTCAGGGGCATATCACCATCGGGTACGGCTTCGATACCTT
CGTGCATGAAGCGTCCGAGCTAACTCTTTGAATCTTGTGTTCTACGCGACAGAAGGTATTACCTGCATTGCAGCTATCAA
CGTCCGACCCAGGCTTCTGGAGCGTCTATGCCTTGTCTGGACAAAGTCTCACGGATGACGATGGGCTATTACTCTTTAGTGCC
AAAGCGCGAGCTGTTGTTCAACGCATAGCAAGCAACCAAGTTTGCAGGTAAGTGAATGGGCTACCCCAAGCTATCAAAAACGGT
TGCGCTTGATCTATATTATCAATATGGGCAGACTGGTAATTTTCAAAAATTTCAACAAGCTATAAATAGCCATGATTGGCCGG
CAGTCATCCATGAACCTAGAAAAGTGAATGGTGTACCGAATGATCCTCTCCAGTTTATTACAAAACGATTGGAAGAGCGAGCC
AAGTATCTGGCAATATCCTTCAACTATGAGCAATGA

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND
PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September 12, 2003 Serial No.: Not Yet Assigned

Page 46 of 118

Customer No.: 21559

47/118
FIG. 30B

Protein sequence: (SEQ ID NO: 155)

MAEAIKRDAMMTKLYFDLLNSPAEAHSSIQKSLSVQAISTTVPILEFPSETVYAYASYINALSIGQRIDPAFTQSLTSAISNL
AGRPIAVSDIYQKIHETTLRTPVEMGVRPNSITFEEYQATINQQAINMVQDMQDGDKEKVEALQANMQFLYQGEINTDFIAR
NELAAGQRAKTVAIVQGHITIGYGFDTFVHEASELNSLNLVGSTRQKVLPALQLSTSDPGFWSVYALLGQSLTDDDGLLLFSA
KARAVVQRIASNQFAGKWNGLPPAIKTVALDLYYQYGQTGNFPKFQQAINSHDWPVIHELNRWNGVPNDPLQFITKRLEERA
KYLAISFNIEQ.

RL055

DNA sequence: (SEQ ID NO: 34)

ATGAACAACACAGTGAGCGAAACGCAACAGATCAATATTTACCAAAATCCGGGGCAGTCTATTTCCGGTCTCTACAAGGGGCT
GGCTAACCAAGTGTCTCTCTGGCCAGCCATTTCCAGAGGTACAGCTTGTGGAGGCTTGGGATATCCCTCTCGTACTCCATCCGG
AGTTTGTGCCTAACCGAGATGTCTCGAAAATCGATAAGGAGTACGGAACGATCCTTGCTGCTGAGTCAGCTCAGGTTATCCTG
CTTCAACTCCAAATGGCTCAAGACAAGGCTAAGGCGTGCGGGGAGGTTACAGCCTTGATCAGTTCTGTCTCTCCTCAATCTCAA
TACCATTAAGAGTCGTATGGTGCTAATTATCTAAACCTGCTGAAACAATCACCGAACCGATACCCGACTAGCGTCGGAGTTG
AGATCATGTGAGTGGCAGTCCGAACCAGGATTCTGGAATCGAGGTCTCTTACGGTGCCAGTCTCGGCCGTCTAACTCAATCA
CAACTTCAGGCGATGAATCTGCCTGCCAGTCTCAAACAGTTGCTCACTCAGGGAATCGGTGTGAAGCTTTCTCAGCCTGAATA
TTGGCCTGCTTACAACAACATAGCCACTGGTATTCTGTTATACAACCGAGTGCGGATAACGTTGGCCTATTGGGCCACGGTTT
AG

Protein sequence: (SEQ ID NO: 156)

MNNTVSETQQINIYQNPQGISISGLYKGLANQCSPPGFPEVQLVEAWDIPLVLHPEFVPPNGDVSKIDKEYGTILAAESAQVIL
LQLQMAQDKAKACGEVNTALISSVSSNLNTIKSRHGANYLNLLKQSPNRYPTSVGVEIMSGGSPNQDSGIEVSYGASLGRLTQS
QLQAMNLPASLKQLLTQGIGVKLSQPEYWPAYNNIATGIRYTTGVAITLAYWATV

RL056

DNA sequence: (SEQ ID NO: 35)

ATGACCCAAGCTGCGAAAATACCAGCAAATGAGTACTCATTGGGGGATGGAAGAGGCTACATCAATATCTGGCCGGAAGGA
TGAGGCTCAGGCATTTCTTATCCATAATGATGGGCCTAATGGGGCTACATGCAGCCTTAAAGGCACTCTTAGAGATAATAAG
GAGTGGTGCAATTCGCCGTATTCCTCTGCTTCATGTTTGCTAAGTATCACCCAGACAGGGCTGCTGTGAGTAAGCGTCAAACGT
GAGGAAAATTCGCCAAGCTGCTCTGCATGGTGCGGTCTTAGAGTTTGGTTTGAAGGAGCCTATAGCGTCCCGCCCAAGGGCTG
CTACTATATGCAAAATAAGGAAAAAACTCGACAAATGTTGGGTATGATTGAGAAAAAAGAGCTTGATGCCGCTCGCGCCTTAT
CAATAAGCTTTTGTGCACTGCGCAACCGAGCTAGCCTATCCTGCCAAGATATACTTGACGAACACACTTGCCATGATCAGT
GCTGAAAAGGGAGAGAATGCTCGCTGTTTGGAGTATGCCCATCGGGTGCAAAAGCAAATTCCTGTAAGAGATGACGGCCAACC
GGCTGAAGACTTGCTCCCGGCGGAGCACGCTTTTCGTATGGAACAACGCGCCAAGGCTGATGCTCTGCTGAGCGATGCACGG
ACGAGAAATAA

Protein sequence: (SEQ ID NO: 157)

MTQAAKIPANEYSLGDGRGYINIWPEKDEAQAFLIHNDGPNGATCSLKGTLRDNKGVVHSPYSSASCLLSITQTGLLSVSVKR
EENSPSCSAWCGPRVWFEGAYSVPKGCYYMQIRKKTRQMLGMIEKKELDAARALSNKLLSDCATELAYPAKIYLTNTLAMIS
AEKGENARCLEYAHRVQKQIPVRDDGQPAEDLLPAEHAFAMEQRAKADALSERCSEK

RL057

DNA sequence: (SEQ ID NO: 36)

GTGCTGGTAGAGCGTTTGCCGACTGATGTTGAATTCGCGGGCGAGCTGAGCCTTGGGCTCGCCGGCCGCTGCCCGCAGCCCCA
GGGTAGCACCTGCTTGTCGACAAAGGCCTCTTTGCGGCCCGGTACGCGCAGAGCTTGATATCCTCGCGCTACCGCGCTGGTG
CTGCTTGATGCTGCTCAGCAAGCCAGCCGCCGCTCTTTTCAGGGTAAGCGTACGGCCAATACACCTTTACTTAGGTTGA

Protein sequence: (SEQ ID NO: 158)

VLVERLPTDVEFAGELSLGLAGRCPPQGSTCLSDKASLRPRYAQSLISSRYRAGAACMLLSKPAAGLFRVSVRPIHLYLG

RL058

DNA sequence: (SEQ ID NO: 37)

ATGGATATTGCGCTGGAGATTTTAGCGCTTGAACAGCTGTTGCTAGAGCCGGAATCGAGAAAGAATGATCGACTGCTTAAACA
GCTGCTTACCGAAGACTTCGTTGAATTTGGAGCTATCGGCAAAAGCTGGACGAAAGCGGAGGTGATCGTGGGACTAAAATCCC
AGACTTGGATCAAAGGACAATCGAGGATTTCAAACCTGCGTGTGCTTGAGATGGTGTGCGGTTAGCAACGTACCGATGCCGT
CATCAAATGCTAATGGCGATGAGTCGTTATCAATGCGTAGCTCTGTTTGGAAAACCTACGAAGATGGTTGGCACATGGTGT
TCACCAAGGCACGAGGGTCTCCGAGTAG

48/118
FIG. 30C

Protein sequence: (SEQ ID NO: 159)

MDIRLEILALEQLLLEPESRKNDRLKQLLTEDFVEFGAIGKSWTKAEVIVGLKSQTWIKRTIEDFKLRVLADGVALATYRCR
HQNANGDESLSMRSSVWKTYEDGWHMVFHQGTRVSE.

RL059

DNA sequence: (SEQ ID NO: 38)

ATGACTTCCTCGCCCAACCTTGACCAGATGACCCCGGAACAGCTTCGTGCCTTGGCGGCACAGGCGTTGCAGTTGCAATCCCA
GGTCGAGGCGATGAGCAGGAAATCCGCAACAATGAAACCCTCATCGAACAGTTCAAGTTCGAAATCGCTCTGCTCAAACGCC
ACAAGTTTGCCAAGCGCAGCGAGCAAATCAGTTTCGGCGCAAGCGAGCTTGCTGGATGACCTGCTCGACACCGACCTTGAAGCT
ATCGAGGCCGAGCTGAAACAACCTCTTCCAGCTTCGCCACAAGCCGAGCCACGGCAATCCCGAAACGTTTCGCCATTGCCGCC
GCAGTTCCTCGCGCACGGTGATTTCGCCACGAACCTGAAAATACCCAATGCGCCTGCGGCTGCCAATTC AACGCATCGGCCGAAG
ACGTGAGCGAGAAGCTGGATTACACGCCGGGCGTGTTTACCGTCGAGCAACATGTGAGGGGCAAATGGGCGCTGCCGTGAGTGC
GAAACCTGATCCAGGCGCCGGTGCCAGCCAGGTTATTGATAAAGGCATCCCGACCGCAGGTTTGTGGCCACAGTGATGGT
GGCCAAAGTTTGCCGATCACTTGCCGCTGTACAGACAGGAAAAATCTTTGGCCGCGCCGGGCTGCCAATTGCCCGCTCGACCC
TGGCGCAGTGCGTCCGACAAAATGGCGTGCGGCTTCAGCCACTGGTCGATGCACTGCGTGAAGCCGTGCTGAACCAGGACGTG
ATCCACGCCGATGAAACACCGGTGCAAATGCTTGACACAGGCGAGAAGAAAACCCACCGGGTCTATGCTGGGCGCTACAGCAC
GACGCCGTTTTTCGGCGCTCAAAGCGGTGGTTTACGACTTCAGCCCAAGCCGTGCCGGAACATGCACGCAACTTCCTAGGCG
ACTGGAATGGCAAGCTGGTCTGCGACGACTTCGCTGGATACAAGGCCGGTTTTGAACAAGGCATCACTGAAATCGGCTGCATG
GCTCATGCTCGCCGCAAGTTCTTCGACCTGCATGTGCTAACAAGCCAACTGGCCGAACAGGCGCTGCACTCAATTGGCGG
TTTGTACGAGGTTGAACGCCAGGCTCGGGACATGAGCAACGAAGACCGTTGGCGAATACGTGAGGAAATGGCGGTACCGATCA
GCAAAACACTGCATGACTGGATGTTGGCCAGCGCGACCTGGTGCCTAACCGGCTCGGCCACAGCTAAAGCCCTCGACTACAGC
CTGAAACGCTGGGGAGCGCTGACGCGCTACCTGGACGATGGGGCTGTGCCATCGACAACAATCAGTGGAGAACAGATACG
GCCGTGGGCGCTCGGACGCTCGAACTGGTTATTTGCCGGATCGCTGCGCAGTGGCAAACGAGCAGCAGCTATCATGAGCCTGA
TCCAGTCCGCTCGCATGAACGGGCATGATCCGTATGCCTACCTGAAGGACGTGCTAACTCGCCTGCCGACGTTACGGTCGAAA
GACATCAGCCAGTTGCTGCCGCATCAGTGGGTACAGATCTAG

Protein sequence: (SEQ ID NO: 160)

MTSSPNLDQMTPEQLRALAAQALQLQSQVEAMSRKIRNNETLIEQKFELALLKRHKFAKRSEQISSAQGSLLDDLLDLDLEA
IEAELKQLLPASPQAEPRQSPKRSPLPPQFPRTVIRHEPENTQCACGCLQRIGEDVSEKLDYTPGVFTVEQHVGRKWACRQC
ETLIQAPVPAQVIDKGIPTAGLLAHVMVAKFADHLPLYRQEKIFGRAGLP IARSTLAQVWGQTVGRLQPLVDALREAVLNQDV
IHADETPVQMLAPGEKKTHRVYVWAYSTT PFSALKAVVYDFS PSRAGEHARNFLGDWNGKLVCDDFAGYKAGFEQGITTEIGCM
AHARRKFFDLHVANKSQLAEQALHSIGGLYEYERQARDMSNEDRWIRIQEMAVPISKTLHDWMLAQRDLPVNGSATAKALDYS
LKRWGALTRYLDDGAVPIDNNQVENQIRPWALGRSNWLFAGSLRSGKRAAAIMSLIQSARMNGHDPYAYLKDVLTRLPTLRSK
DISQLLPHQWVQI.

RL060

DNA sequence: (SEQ ID NO: 39)

ATGATCCGCATCGATGCGATCTGGCTAGCCACCGAACCGATGGACATGCGCGCCGGCACCGAGACGGCATTAGCCCGGTAAT
TGCGGTGTTTCGGTGCGGCGAAGCCGCACTGCGCTTATCTGTTCCGCAATCGCCGGGCTAACCGAATGAAAGTGCTGGTGACG
ATGGCGTGGGCATCTGGCTTCCCGCGCGTCTGACTGAACCAAGGCAAGTTCCACTGGCCCGGCATTGCCATGGCTGCGAGGTC
GAATCGACAGCGAACAACCTCCAGGCCTTGGTGCTGGGCTGCCGTGGCAGCGCGTCCGCACAGGCGGTGTGATCAGCATGCT
GTAA

Protein sequence: (SEQ ID NO: 161)

MIRIDAIWLATEPMDMRAGTETALARVIAVFGAAKPHCAYLFANRRANRMKVLVHDGVGIWLAARRLNQKGFHWPGIRHGCEV
ELDSEQLQALVGLPLWQVRVGTGGVISML.

RL061

DNA sequence: (SEQ ID NO: 40)

ATGCGCCAACGAAGCTCTTACCCGAAACCGTTCAAAGCCCAGGTCGTTGAGGAATGCCTGCAACCTGGGGCAACGGTGTCAG
TGTCGCCATCAGCCACGGCATCAACGCCAATGTATCGGCAATGGCTGACGCTTTATCGAGACCAGCCGTACCAGCCTCGT
TACCAGCCTTTGTCCCGTGAAGGCCACCCCTAAACGGCCAGCGAAACGTCAGTGCTCATTGAAGTCCCATGGCCGGGCAA
ATGATCACGGTGAAATAG

Protein sequence: (SEQ ID NO: 162)

MRQRSSYPKPFKAQVQVECLQPGATVSSVAISHGINANVIRKWLTLYRDQVPVPSLPFAFVPLKATPKRPAETSVLIELPMAGQ
MITVK.

49/118
FIG. 30D

RL062

DNA sequence: (SEQ ID NO: 41)

ATGGCTTTATCTCTTATTCGTAGTCTCACTGCGTCCGCCTCAGGAAACATCTCGGCGTTGAAACGCGATGCCAAACGCTTGCA
GAAGAACTCCTTTCTTGTGTTTGAACAGAATATCCACTCAAGGTTTGCCAAAATGCGGTAGCAGTTTCTCGGGCTTCCGCT
CACTGCTGATGTCGATAAACTGGAGCAGCATTGGCATGAATAGAAGCGCTCCATTCTGGGTGATCCGTGGCCGCAACGAT
ACACACCAGGGGGTACTGGAAGCGCTATATGTTTACACCTTGAATATACCGAGAATGGCCCCGTCGTTTTACTGGAACCC
AAAGCACTCTATACTTCAGCCTTAGTCCTTTTCTTGAGCAAATGAGCTTTAAGAACTACCCGGAATAATCCTCATCGAAA
CAAAAGAGACCTCAATCCAAACAACCCATATATTCGACGCAATAGAAAAATTAGAAGTCGAAGAACTCTAAATAAAATTCGA
TTTCTTGACTTGCGAGACCGAAACCTTCCCGTTTCGCTTAGTACCGAGGCTCGTTGCTGGATCGAGTCAATTGTCAGTTTATT
GCCAAACGACATCCAAGAGGAAATACGTAATAAGGATGGTCAACTCACTTAGAGATCAGTGCATATGAGCATGCAAAGTCTC
GTAATCAAGTATTTGGCTCCTCCAAC'TTCCCTTGCGTCCCTTCTCTCCATAAAGTCAGCGATCTATCAACTCATTTAGGC
GCATACCTCCCTTATGGATGCAGCCATCCTCCTTGCGGAAATATCTAAAGTTGATATACGCCGACCTCCTCTCGAAAAAG
CTCAGAGGAAACCTTACTTTATCTCATAAAAAATTAGAGAATCGACAGTTCCACACAGGCATTTTCATGTGAGCATGAGAGTC
GATGGCGGCCGTATGTCGTACTCTTCTCCAGGAATGATCCGGCTAGCGAGGTACTAGCAGGAGTTATACACTCGTACTTTTCT
TGGAAGCAAGATAGAGACCATCGCTCACCCACCTTTATGTTTCAGATGGAGCAGTTCCCTATGCTCCCAAGCTTCTAGGTTT
AGGCGGCCATACGGTCATTGCAAATGGAATCACTGAAATTCGCCAGCGGGATGGTCTTGGGGAGTTCTATGGCTACAAGAACT
CACTTAAAGTCAGCTCCTTATCTAACGGAATACAGTTCATGGGTAAGCATGTATCACTAAAGTAA

Protein sequence: (SEQ ID NO: 163)

MALSLIRSLTASARNISALKRDAKRLQKNSFLVFGTEYPLKVCQNAVAVSRGFRSLADVDKLEQIHGMNRSAPFWVIRGRND
THQGVLEALYCLDLEYTENGPPVFTGNPKHSILPALVLFLEQMSFKKLPLGLILLETKETS IQTTHIFDAIEKLEVEETLNKFR
FLDLDRNLPLVSLSTEARCWIESIVSLLPNDIQEEIRNKGWSTHLEISAYEHAKSRNQVFGSSNFPVFPFLSIKSAIYQLISG
AYPPLWMQPSSSGEISKVDIRRPLEKSSEETLLYLIKLENRQFHTGISCEHESRWRPYVVLFSRNDPASEVLAVIHSYFS
WKQDRDHRSPPLYVSDGAVPYAPKLLGLGGHTVIANGITEIPDGDGLGEFYGYKNSLKVSSLSNGIQFMGKHVSLK.

RL063

DNA sequence: (SEQ ID NO: 42)

ATGAACGCTCTGACCCAACCGGCCGCCCTCGCCGCTCCACCTGAACATCAACCTGACCGACTTCATCGACGAGTTCGGCGA
CGAGCTCCTGGAGTCGCTCAATCGCTCCAACCCCCCGGTCTATACCGGCTCCGTCAACGCTCACCGCCAGTTGGTGATGGACC
GACTCAAGCGCAAGCCTTTCGCGGCCACGGCCGAGGTCTGCCAGGCCATCACCGCCTGCTGCTGGACCGTAACGAGCAGGCC
GGAATCATCAACGCCGAGATGGGCACCGGGAAACCATGATGGCCATCGCTGTGCGCAGCGGTATGCACGCCGCCGGCTATCG
CCGGACCTGGTCTGCTCTCCGCCGACCTGGTCTACAAGTGGCGCCGCGAGATCCTGGAGACCATCCAGCCGCCCGCTCT
GGGTACTCAATGGCCACAGATACTCTACTCAAGCTGCTCAAGCTGCGAGATCAGATGGGCGACGCCTACGACGGGCGCCAGGAG
TTCTTCATCCTCGGCCGCTGCGGATGCGGATGGGTTTCCACTGGCGGCTCGCTGCTGGAAGAAACCGCCGCCGGCGGCCA
ACTGCTCGCTGCGTGCCCGGATTGCGGACAGGTCTCTCAGGACCTTGAAGGCAACCTGGTCAAGGTGGAGGAGTTTCGAGCGTG
GTGACCGTTCGACGTACCTGTTCTCTGCGGTGGGGCGCTCTGGACGCTGATCCGACCAGGCAAGCCCGACGGCGGCAACCGG
CGCGCAACGATTCTCAAGTCGATGTGCCGATACCAACCATCGCCCCGTGAGGGCGGAGCGCCTGCTGAACGACTTCGGCGA
GGACTTCCTGGCCACGATGTTGGTGGACAACGTCTCGGAGTTCATCAACCTGATGGACGCCAAGGGCAACTTCGTCTTCAGCG
ATCGGCAGGCCAAACGCATGGAGCGATCGATGGCAAACATCGAGTTCGGCTTCGGTGAAGGCGGCTACCAACCGACCGAGTTC
ATCAAGCGCTACCTACCTGATGGCTACTTCGACCTGCTGGTGTGGACGAGGGACATGAGTACAAGAACAGCGGCTCGGCCCA
GGGCCAGGCCATGGGCGTTCTCGCAGCCAAGGCACGGAAACCGTGTCTGCTGACCGGAACGCTCATGGGCGGCTACGCCGACG
ATCTGTCTATCTCCTGTTCCGCATCTCACCAGCGCATGATCGAGGACGGCTATCGGCCCAACGCCGCGGCGGAGCATGGCT
CCCGCAGCCATGTGTTTCATGCGCGACACCGTGTGCTCAAGGATATCTACCCGAGCGCGACGGTGATTTCGCACAAGACAGC
GCGGGGCAAGAAGCTCTCGGTACGCACGGTGAAGGCTCCCGGCTTCGGCCCAAGGGCATCCACCGCTTCGTATTGCCGTTCA
CCGTGTTCTCTGAAGCTCAAGGATATTGGTGGCAACGTACTGCCGACTACCAGGAGGAGTTTCATCGAGTGGCCATGGCGCCT
GAGCAGGCTTCGGCCTATCAGCGCTTGGCGGCCACGCTGACAGCGGAGCTCCGCCAGGCTCTGGCGCGACGAGATACCACGCT
CCTGGGCGTGGTCTCAACGTGCTGCTGGCTTGGCCGAGTCTGTTTCCGACCGGAGATCGTCAAGCATCCGCGAACCCGGG
ACACACTGGCCTTCGTGCCAGCGATCTTCGGTGACGAGCAGTTGATACCCAAGGAGCAGGTGCTGGTGGACCTCTGCTTCGAG
GAGAAAGCGAAGGGCCGAAGGTTCTGGCATACCGCTCTACAGCGGGACGCGACACACGCTCAGGCTGAAGAAAGTGCT
CGAGCAATCCGGGCTGAAGGTGGCAGTGCTACGTGCTTCGGTGCATACCGCTCGACGCGAGGATTGGATCCTCGACAGGTGCG
ATCGCGGCATCGATGTGCTGATACCAACCCGGAGCTGGTGAAGACCGGGCTGGACTTGCTCGACTTCCCGACCATCGCGTTC
CTGCAAACGGGCTACAACGTGTACACCCTGCAGCAGGCCGCGCGCGGCTCGTGGCGGATCGGGCAGAAAGCACCCGGTGAGGGT
GGTGTCTTCGGCTACGCCGCGAGCTCGCAGATCACCTGCTTACAGCTGATGGCCAAGAAGATCGCTGTGGCTCAGAGCACGT
CGGAGACGTTCCCGAGTCAGGTCTCGACTCGTTGAACCAGGATGGGGATTCTGTGGAGATGGCGTTGGCACGACAACCTCATC
GCAGCATGA

50/118
FIG. 30E

Protein sequence: (SEQ ID NO: 164)

MNALTQPAALAASHNLNLTDFIDEFGDELLES LNRSNPPVYTGSVNAHRQLVMDRLKRKPF AAQAEVVQAITALLDRNEQA
GIINAEMGTGKTMMIAVAAMVHAAGYRRTL VVSPPHLVYKWRREILETI PAARVWVLNPGDPTLLKLLKLRDQMGDAYDGRQE
FFILGRVVRMRMGFWRLACWKKRAAGGQLLAACPD CGQVLEDELEGNLVTVEEFERGD RRRRTCSSCRGALWTLIRPGKPDGGR
RATILKSMCRIPTIGPVRAERLLNDFGEDFLATMLVDNVSEFINLMDAKGNFVFSRQAKRMERSMANIEFGFEGGGYQPTF
IKRYLPDGYFDLLVLDEGHEYKNSGSAQGGAMGVLA AKARKTVLLTGTLMGGYADDLFYLLFRI L T QRMIEDGYRPNARGSMA
PAAMSFMRDHGVLKDIYTERDGDGSHKTAR GK KLSVRTVKAPGFGPKGIHRFVLPFTVFLKLDIGGNVLPDYQEEFIDVPMAP
EQASAYQRLAATLTAE LRQALARRDTTLLGVVLNVLLAWPDCCFRPEIVKHPRT RDTLAFVPAIFGDEQLIPKEQVLVDLCFE
EKAKGRKVLAYTVYSGTRDTSRLKKVLEQSG LKVAVLRASVDTARREDWILDQVDRGIDVLITNPELVKTGLD LLDFFPTIAF
LQTGYNVYTLQQAARRSWRIGQKHPVRVVF FGYAGSSQITCLQLMAKKI AVAQSTSGDVPESGLDSL NQDGD SVMELARQLI
AA

RL064

DNA sequence: (SEQ ID NO: 43)

ATGGCCCTCATGTTCCCGCGCTTGGCGCGCAACTTTGCACGCAACGGCTACTTCCCTACCGATGAGGTCA C C C T C G A A C G C G C
TCTGCAGGCCCTCACTCTTGC C C C C G T C G G G A A G G A T G A G G A T C T G T G A C C C C T G C G C C G G T G A G G G T G T T G C C C T G G C T G A G G
CAGCACACACCCTCGGCCGCGATCAGGTCCAAGCCCTCGCTGTGAGTACGACCGGAGCGCGCCGACCATGCCCGAGGATTG
CTTGACCGAGTGCTGCACAGTGACCTTTTCGACACCATGATCAGCAGGCAGTCGTTCCGGA CTGCTCTGGCTCAACCCGCCTTA
TGGCGACCTGGTGGCGGACCACTCCGGTGCCTGCGAGTACCAGGGCAGCGGCCCGCGCGCTCTGGAGAAAGCGTTCTACCAGC
GCTGCCTGCCGTTGCTGCAGTACGGCGGCGTCATGGTTCTGATGTTCCTCACTACGTCTTGGACGATGAGCTGACTGGCTGG
TTGAGCAACCACTTCACCGGCCTGCGCATCTACGAGCCGCGGATCCTACCTCAAGCAGGTGGTGATCTTCGGCATCCGGGT
CCGTGCGCAGGACCTGGCCCGGGCGGACGCCAATCAGGTGAGGTCTCGCTGCAGGCGATCGGAGCGGGCCAGGAAAAGGCCG
AGGAAATTCCAGCGGCTTGGCCGTGGGAACCTACGTGGTTCTGCCGGCCACCAGCGAGCTGGAGCACTTCTACCGAGTAACC
CTGGAGCCGGAGCAGTTCGCCGGTGAAATCCAGCGGCTGCGAGGTCTCTGGCCTGACTTCAACCTGCACTTCGCGCAAGCGGG
GCTGCAGCCGCGCCACAGTCCGCGAGCTGTCTCGCTGGCACCTGGCCCTGGCCCTGGCCGCGCGCGCATATCTGGCGTCG
TGCGATCGAAGTCGGGCCGGATCCTGGTCTGTGAAGGGTGACACCTACAAGGACAAGGTCCGCAAGACCGAATTACCGAGGAC
GACGACGGCAACATCACGAGGTGAGGATCCTCACCAGCCGTTTCACTCCCGATCATCCGGGCATGGGAAATGACACCCTCTC
GGTCAATCAGGGCCGCGTGCTGACCATCAGCTCCTCGGCCGCGACCAAGGAGCTGAAGAGCCCCAACCTGAGCCGCGCC
CCGACCCGCGACCGCTGCTGATCAGCCCTGGCCGGGTCTGAATGACCGCAGCCGTGAGCCACCTGGTGGAAACCGGTCAACTC
AACCACGCGCCTTTGCTGAAACGCCATCTGGCGGGAGATTGGGGAACGCTGGACCAGGAAGACTGGAACACCAACCAGAGAGC
CCTGAAGTTCCGCGATCGGCTGCTGTCTCCTACGACATCGACGCCGGCGACGAATCCAGGCTCTGGATCATCACTGAGGCAG
ACCGCAGCTCAACCACGCTTTTGCTCCCTAGCGATTACTGA

Protein sequence: (SEQ ID NO: 165)

MALMFRLARNFARNGYFPTDEVTLERLALQALT LAPSGRMRICDP CAGEGVALAEAAHTLGRDQVQALAVEYDRERADHARGL
LDRVLHSDLFDTMISRQSFGLLWLNPPYGD L VADHSGASQYQSGRRRLEKAFYQRCPLLLQYGGVMVLIVPHYVLDDDEL T GW
LSNHFTGLRIYAAADPTFKQVVI FGIRVRRQDLARADANQVRSRLQAIAGQEKAEIIPAAWPWEPYVVL PATSELEH FYRVT
LEPEQFAGEIQRLRGLWPDFNLHFAQAGLQPRPPVRELSRWHLALALAAAGISGVVRSKSGRI LVVKGDYTKDKVRKTEFTED
DDGNITEVRILTD RFIPIIRAWEMTPSSVNQGRVLTIS SAAATTEEAEEPQPEPAPAPAPLLISPGRVMTAAVSHLVETGQL
NPAPLLKRHLAGDWGTL DQEDWNTNQRLKFGDRLLSSYDIDAGDESRLWIITEADRSSTTLLLP S D Y .

RL065

DNA sequence: (SEQ ID NO: 44)

ATGCCCAGTCCCACCCCGCTCTACCAGATCGAAGAGTGTCCAGACCTGTACGTCGACGCTGCGTGTGCGACGAGCAGTGCAA
CCTGGTCTTTCTTTTCGGCCTGGGGCCGCGACACCGTGACACAAGAGTTCTTGCCAGGCTGACGCTGGGCCGGGAAGAAAATG
GCATCGACCATTTCACATCATCGTGGACGGCCGCGCTTACCTGTCTTCCCAAACAGGATCTCTGGAGAAACGCACCACC
CGTCAGTTCGCGGCGACGTTGTTGCGCAGCCTGCTCAATCTTTGGCTGTTTCGATCGGCGCGCCTCGGCGCCCGACCGAGGCAA
TCACCTCGCCTTCGCACTCCTGCAGCGCGATGAGGATCCACACCAGAGGCTCTGGCCGCTGGTGATGAAACCTGTCCGCTCC
CCCTCTGCAGCACTGGCGCGAGCCGGTGATGGAGGTTCTACCCAGCACCAGATGTTGACGGCCCTACCCGGGACGATCGGC
AACGCTCTGCGCCTGGCGACTCGCCCTGCGGGTCGACGTGCTTGAGCCCAACCTCGGTGAGGTAATCCGCGAAAGCATTCTTAC
CACCGATGCTCAGGCGCAAGCCTGA

Protein sequence: (SEQ ID NO: 166)

MPSPTPLYQIEECPDLYVDACVCDEQCNLVFLSAWGRDVTVTQEFARLTLGREENGIDHFHI IVDGRRLLPVFPNQDLLEKRTT
RQFRGTLFGSLLNLWFD RRRASAPDRGNHLAFALLORDEDPHQRLWPLVPMETCPLPLLQHWREPVMVELTQHQM L TALPGTIG
NVCAWRLALRVDVLEPTLGEVIRE SILTTDAQAQA

51/118
FIG. 30F

RL066

DNA sequence: (SEQ ID NO: 45)

ATGAATCCATTGTTACCAACCTCACCCAGGAAACCTCGCTACCTCGAGGACCAACTGTCCAACAACGACGTGCGCCGCGA
CGACGAGCTCATCGACTTGTTCATCGAGGAGCTGTGCTGACCTTGGAGCAGGCGGAAGCGGCTGTGCGCTACGCGATCAGT
ACCTCTGCCAGGTCTTCTGATCGGCCAAGGGCCGCTGCACCAAGCCGATGGACTCAGCTTCGACCCTCACACCAAGAGCGTT
CGGTAG

Protein sequence: (SEQ ID NO: 167)

MNPLFTNLQTETLAYLEDQLSNNDVAGDDELIDLFIIEELSLTLEQAEAAVALRDQYLCQVFLIGQGPHLQADGLSFDPHTKSV
R

RL067

DNA sequence: (SEQ ID NO: 46)

ATGGGATGGCTTTTCTCACATCAGACGAAGGAAGACCTGCTGCGTGAGCTGCTGGCCCCAACCAGTACCTTCGAGGCAGCAC
CGAGGTGCTGGCACACGAGTCTCCGGCAATGAACCTTTGGACTGTGCTAAAACGAACCTTTTACCTTGCCGGATTCTATTTCC
GCAAGCCGGCCGGTCACTCGATCACCATGATCGAGCTGCACTTGTGCTGGACTGCTCGGCCGGGCAATGGGGCTACAAGACCATT
CCGGAAGCGCCGGCCGTTCTACTACGGCTGTCCGCTGGAGTTCTTGACCTGGCTCACGATGAGATCAACCAGGAATGGCG
TAAACGCCTGACGCACGAACACCAAGCCTGA

Protein sequence: (SEQ ID NO: 168)

MGWLFHQTKEDLLRELLAPTSTFAGSTEVLAAVSGNELWTVVKRTFHLAGFYFGKPAGHSITMIELHLLDCSAGQWGYKTI
PESAGPFYYGCPLEFLDLAHEINQEWKRRLTHEHQA

RL068

DNA sequence: (SEQ ID NO: 47)

ATGAAATCGATCTACAACACCCAGGCTTCAGCGAGGAGTTGTTGCTGGTTTGGCCTCGCTGCGCGAGGTGCGACTGGACAA
TCTGGCTGACCAAGTTCCGCGCGGCAGTGTTCGACCGATCCGTGCTGACCAGGCCATCATCGACTGCGTGAGCGGGTGAAGA
CCCCCTCGCCGGAGCATGCGGCCGACAACGAGCCCTGGTTGTACTGCGACTGGCAGGCCAGGCAAACAGCTTACCGGCTCCTC
CAGCGCCTTGAGCGCGCAACACGCTGA

Protein sequence: (SEQ ID NO: 169)

MKSIYNTPGFSEELLVCASLREVGLDNLADQFRAAVFDRSVVDQAI IALRERVKTPSPEHAADNEPWLYCDWQARQTAYRLL
QRLERATR

RL069

DNA sequence: (SEQ ID NO: 48)

ATCCCCCTACCACGATCCCGCTTCGGCGGGATCATCCTTTTCGAGGTTCATACCATGATCACAGTTCCTCGGACAGTTGGCCAT
TCGAACCATCAACGGTCCGTATGGCGAGTTCAATGTGGGAAACTCTGGACTTCGATCGGGGAGTTTCATCATCAAGGATGCCT
TCCTGGATCAACACACCGAAGGCAAGTACCGCGGTGATTTCTGTCATCGCCAATATCCGCCCCCACCCTACTCCGCCGGCGGT
CGGCTAGTCATCGAGATCCGCGCCATAGTGGACAGCATGACGCTGAACGATATGGACAGCCTCAGCGACGAGGAGGTAGAGCG
TCTTTCCGGCAATGAGGTGGATCCGCTCGACGAAGTGCCCGAGATCCAGCTCCCCACAGTAGTACCGGCGATACCACCAAAGT
CGCCGTACCCCCAGAAGTCGAAGCCTCTGTGCCTCGCTGCAACAGGGACGCGCCTTTCGGTATGGACACTCCGGCTCCTGCA
GAGCAGGCCGCTCTCTGACACAGACGCGGATGCAGAACTGTTCGGGACGGTCTGGCCGCTAGGCGAAATCGTCAAGCTGGA
CACCACGGTGCACCGCAAGCGACTACGCCAACAGTGCCTGCGACTCGGCGCGCTGGGCTATGAGCTCGACTTCAAACAACAGG
TGTGGACCCGCAAGGAGCCGCATGA

Protein sequence: (SEQ ID NO: 170)

IPSPRSRFGGIILFAGHTMITVPGQLAIRTINGRYGEFNVGKLWTSIGEFI IKDAFLDQHTGKYRGDFVIANIRPHYSAGG
RLVIEIRAIIVDSMTLNDMDLSLDEEVERLSGNEVDPLDEVPEIQLPTVVPAP

RL070

DNA sequence: (SEQ ID NO: 49)

ATGACCTCTCTCAACAACCACTCCAGCGCAGGTACACTGCTGCGTACCTCAAACCTCCCGATCGTTCTACCAACGCGGCCTG
GCTGCGCCTGGTCTATCTCGCAACCTGCCAGGGTCGACGAGATGGGACCCGGCTGGCCAGTGTGCTTCAAACCGCCTGGC
AGGAGCTTTCTCTCCAGCCGACCGCGAAGCACATCCAATTCCACCTGTACCACAAGGAGGAAGAGGGGAGGACCCGCGCTC
GCGTGTGCTGGTTCTCTCGATAGTCGAGCCGTCGATGAGCCTTCTACCTGCGCATCGAGTTGCAGGAAGAGTGCCTCGCCGA
ACACCCGGTTACCGAGTAG

52/118
FIG. 30G

Protein sequence: (SEQ ID NO: 171)

MTSLNNHSSAGHTAAYLKLPVLTNAAWLRLVYLANPARVDEMGTSLASVVQTAWQELSLQPTAKHIQFHLYHKEEEGQDRAL
ALLVLSIVEPSDEPSYLRIELQEELAEHPVTE . PKSPSPQKSKPLCLAATRDAPFGMDTPAPAEQAASLDTDADAELFGTVW
PLGEIVKLDTTVDRKRLRQQCVRLGALGYELDFKQQVWTRKEAA

RL071

DNA sequence: (SEQ ID NO: 50)

ATGACTCAACTCAACCCGTTTATTTCGCGGCTATGAGAGTTTCCGCATCGAGCGAAACCTGCAGATCACTGACGAAGGCAACAA
TCTACCGTGCTACCGCGCTCTGCATGAAACCCAGCAGCACCTCCAGACGAATATTTTCAGTGCGAGCTGTGCTACTTCAATA
ACGATTTCGCGGTGGTAGTCCAAGAGTTAGACGATGAAAGAGTTGAAAAATGCCCTCACCAGGAATAGTGAGAAACGTACTT
TACAGCATCTACGGTGAGCAGGACGGCAGAAAAAAGCTTATCGGAGATCAATACTCACTGACCGAAGCCGAGAGTGTCTGTTG
ATACCTTTTCGTTTCGGCGCGGTTATAACCCCTGCTGGGAGATCAGAAAAACACATCTACCCATCAGCGCGTGGAATAGCCTCT
ACGAAAGGTTCTCGACCAAGATGCCAATCCGCTTGCCCTCGGTGTTGGTATCGCTCTTCTGGTGTAAACGAGCACGGTGCCGTG
GGCTTTTCGCTTGACAAACACCCCTTGGACGGATGAGTGTCTGGAGATCCTGGAGATGACCGCAGCCGCTCTTCGACAAGAACA
GCTTGCCTTCGGCCTCGACGAACACCTTGTGATCTGCTTACCTCGCGGACAAGCAGACATTCGGCTCCTGGTACTTGATC
CATTTCGCGCCACGCTCAAGGGCCTGCCGCTTTATGACGATTGA

Protein sequence (SEQ ID NO: 172)

MTQLNPFIRGYESFRIERNLQITDEGNLPCYRALHETQQLPDEYFQCELCYFNNDFAVVVQELDDERVEKCPHQGIVRNVL
YSIYGEQDGRKKLIGDQYSLTEASVVRYLSFGGGYNPCWEIRKTHLPISAWNSLYERFSTKMPIRLPSVLVSLFWCNEHGAV
GFR LHNTPTWTDCELEILEMTAAALRQEQLAFGLDEHLVDLLHLAQADIRLLVLDFFAPTLLKGLPLYDD .

RL072

DNA sequence: (SEQ ID NO: 51)

ATGGGACTGGTGTTCCTACCGAAAGGAGAATCACCATGCAATACGGAAAGCTGGCGCTCGCCCATCTCAGCCTGGAAGTCC
GTTGCAAGTACTTATGAATAAGAACCGTGCTTACTACATCGGCACTTCTGACGAAGAAGGACCAGCCTCGCGCGAGTCCGTTG
AATATTACCCCTCAGCGAACTTGCCCAACAGGCATTAGACCACGGCACTTGGACGCAACTGGAATATTA

Protein sequence: (SEQ ID NO: 173)

MGLVFPPTERRITMQYGLALAHLSLELPLQVLMNKNRAYYIGTSDEEGPASRESVEYYPSRELAQQALDHGTWTQLEY .

RL073

DNA sequence: (SEQ ID NO: 52)

ATGGGAAATGTTTGGCGATTATGCCAGGGCAGATACCTGGGCATTGTTGTTGGCCAGGAACAGCCAGGCGAAGTTGCAGAACT
GACTGCTGAGCAGCAGCTCGTCTCGACGTCGCTGAGGCTAACTCCTCAACTTCGGGCAGGGCGGGCAGTTCTACGATTGG
ATGTTGCTCATGATGATCTCCAGATAATGGAGAACACCACGCCCTGGGGGAGATGGTGCCTCCGATGGGTATGCGATGAA
GAGTGGCGCATAGCGTAG

Protein sequence: (SEQ ID NO: 174)

MGNVWRLCQGRYLGIVVGQEPGEVAELTAEQQLVLDVAEANLLNFRQGGQFYDLDDVAHDDLQIMENTTPWGEMVPPGWVCDE
EWRIA .

RL074

DNA sequence: (SEQ ID NO: 53)

CTGACGGGCAAGGTGTTTCTCCGCTTTTCGCTTACGAAACTGGAGAATCATCATGAGCAACAACACCCAAGCCCAAGAAGCCAA
GTATTTTCGACCTGCACACCACCGGTATCGGCTACCTCAATCGCATCCGCGAGGTACCGATCCGCCGAGGTGAACCATTCCTCG
CCGTAACCGTCGAGCCCTCCATGGCGCGGAGACAGCGTGGAATACTCCTACATCGACTGCAAAGTGGTCCGCGCCAGGCT
GAAAAGCTTGTCCGCCGTTGCAAGGAAGCAGTCGAGGCCAAGAAGAAGGTTCTGATTTCTTCCGTATCGGCGATATCTGGGC
GGATCCCTTCATCCACCAGAAAGGCGAGAAACAAGGCAAGCCGACGCAAGCCTCAAAGGCCGGCTGCTCTTCATCTCCTGGA
TCAAAGTGGATGGCACCACCGTCTACGATGCGAAGGAAGAAGCTGAAAAAGCCAGCAAGGCAAAGGCGAACCTCAAGGTGAG
CCCGCAGCCCCCGCTGAGCAGCTGAACAAGCCGCTGCTTGA

Protein sequence: (SEQ ID NO: 175)

LTGKVFLRFRLRNWRIIMSNNTQAQEAKEYFDLHTTGIGYLNRIREVPIRRGEPFLAVTVAALHGAADSVEYSYIDCKVGAQA
EKLVRCKEAVEAKKVLISFRIGDIWADPFHQKGEKQKPDASLKGRLLFISWIKVDGTTVYDAKEEAKEAQQKGEPQGE
PAAPAEHAEQAAA .

53/118
FIG. 30H

RL075

DNA sequence: (SEQ ID NO: 54)

ATGTCCAAGCAATCCACCAGCTTCGAAATCGGCTTTGCCCTCGGCAGTGTGTGCGTGAGTTCGCGAGAGCGCTCAGTCGCCC
TCCGGTCGTAGTGCAAGCACAAAGCGCCGTTGCGTTGAGAGTCCAGCGCATCGATCCTGCCTTCTGGCCGGCCCCGACCGCTG
GCGAGCTAGAACACATCAGCGACATCCAGCCATCGTCCGGCTGAAGAAGGTCAACCTGAATGACTGGTATCTAGCCAATACG
CGCGAGGTGCAAAAGCCCAAGCGCGCACGCAAACCAAGCCGCCAAGGCGACCGCCAAAGCTGAAACGCCAGTCAGGAAGGA
GCTCAAGATGGGTTCCTCTGACCATTTGATTGCACCCAACCTCCGAAAGCGAAATGGGGAGGCCCTCTCCAGTTAGAGTCCC
TGAACGATCATGAGATTGCTCTTTTGCAGCACCTCTGGTAGCGCAGTCTCTTGGGAACCTCCATCGGCGTACTCAGGAGCAA
TACCAACAACGCTGGCAGGACTACTTGTCCACCATGACGGATGAACAAGTAGCTGCTCTCGGCCGCTAA

Protein sequence: (SEQ ID NO: 176)

MSKQSTSFEIGFALGSVVREFRRALSRPPVVQAQAPVALRVQRIDPAFLAGPTAGELEHISDIPAIVRLKKVNLNDWYLANT
REVQPKPRARKPKPAKATAKAETPVRKELKMGSLDHLIAPNSESEMGRPPLQLESLNDHEIALLPAPPGSAVSWELHRRTQEQ
YQQRWQDYLSTMTDEQVAALGR.

RL076

DNA sequence: (SEQ ID NO: 55)

ATGGTGTTCCTCCTGCAGGTTGAGGGTGCGGAGAAAACACTGGCCCTGGCGGGGAAGTGGATTCCCCGCTGGGTGCGGAAGG
GAGCTTCTATCGACCGAGGCCGACCGACCGCTACCAGAAGCTATGCGGTCTTGGGTTGGATCAACACGGTGGGCTGTGCTG
CAGCATTTTCGGATCCGAGCTGCATGGGGCATGTGCTGACAACGTGAGCAGATCACGCGTTCATCATCGAAGCGGGGGCGA
AAGTGTCAAGGTCAAGCAGGAGGGGAGCGGATGCAGCGGGCGGAGAGCGAGGGCGGAAGAGCGCGGCTGGTAGAAACCTGT
CAAAGGTTTCCCCAGCCGTGTCTGGAAGGGGAGTCAAGTGAGCCACCTGTGGTTGAATCGTCGATCCCTGGGCATTGATCGTC
TCGATCCCATCACCCGGCCATTATCGTGGCTTGGCCAGCAAACAGTAGGCACGATCCGCGTACAAAGGGAGCCCTGCGTATC
ACCGCGCGGCCACCGGCAGGGAGAAGGATCCCGATGGGTAGCTGATAGTCTTGGAGCAGGAGCATCAGGCTACCCATGGAGA
GGGGAAGAGGGGCCGTAACACACAGTACGACCTTAAATCGAGGAAACACCGAACCTCTTGA

Protein sequence: (SEQ ID NO: 177)

MVFLQVEGAEKTLALAGKWI PRWVAEGSFYRPRPTDRATRSYAVLGWINTVGCFAAFRIAAWGHVADNVSRSRVHHRSGGR
KCQGGAGGADAAGGERGRKSAAGRNPVKGFPSRVWKGQVSHLWLNRRSLGIDRLDPITRPLSWLQQTVGTHPRTKGALRI
TGGPPAGRRIPMGSILVLEQEHQATHGEGKRRGRNTSTTLKSRKHRTS.

RL077

DNA sequence: (SEQ ID NO: 56)

ATGCCGCTGATGTGGATCGTCTGGTGCTCGCGCTCATCACGGGACCTGGCTGAGTGTAACAAGCGACCACGCGACCTCGAG
CGCCGAACCTGGCCGAGGTGACACCCCTGGCCAGGAGCTTGCTGCTCTTCCGGTCCAGTCTGGCGGAGTACGCACACGCCAAC
CCGGTTTACCGGTTCCGCCGCGGACTCCGCTCTTGGTTTACCGGCCTGGTTCGCAAGCCAGCGCGCTTCAGGGCTACATC
GCCGCCGCCACAGCTACGCCCTTCATCGCCTCGCCGCCGCGGGGCTGGCGCGGCCGTGGATGCTGGTACGGAATCCGACCT
GGTTGGCGTCAGGCGCAACGGCCAGTTAGTCACGCGCCGCCTCGGAGCCACTGTATTGCGCTCCCTACGCCCATCCCCGAGG
GCGCGGTGGTCCGGTCAAATAA

Protein sequence: (SEQ ID NO: 178)

MPLMWIVLVLALITGTWLSVQSDHATSSAELAEVDTLARSLLLFRSSLAEYAHANPGFTGSPADSALGLPAWFRKPARLQGYI
AAGTSYAFIASPPAGLAAAVDAGTESDLVGVRNRQLVTRRLGATVIALPTPIPEGAVVAVK.

RL078

DNA sequence: (SEQ ID NO: 57)

ATGAGGAGTACGCGCAGCAGTGGATTTCATCTCGATCGAACTGATGATCGCCCTCGTCTGATCGCCATCGCGACCGCCGGTGG
CATATCGGTCTGATGAGCTACCTGGACGGCTTGGACGAGCAGCACGCGGCCAGCAGCAACAGCAGGTGGCCAAGGCAGCGG
AGAAGTACCTGAAGGACAACTTCAGCACGGTCTTGCCAGCGCCGGCCACGCCCCGGCGGTGATCACCGTCCCGATGCTG
CGCAACACCCGTTACCTGCCCCGAGGCTTCCGCGACACCAACTCTACGGCCAGCAATACCAGGTCTTGGCCCGCAAGCCGGC
GGCCAACAGCTCGAAACGCTGATCGTGACCACGGGTGGACAGGTAGCTTCCGAACCTCTCGATCCGCCGGATCGCGCAGCTCA
TGGGAGCCACCGGGGGCTACATCTCGAAAACCAACACAGTATCGCCCAGGGCGCGCCTGGCAGGTGGCCTTAAGCAATTTT
GGTAGCGCTCCCGGCGCTGGACATCTGGCGACGGCGCTGTTCTCCAGGACGGCGCCATCGCCAACGAGTACCTCTACCGCAA
TGCCGTCCCGGTTCATCTGAACCTCAACCGGATGAATACCACGCTGGACATGGGAGGCAACAATATCGCCGAGCCGGGGCGA
TCACGGCCAGCGGCAACATCAACACAGCGCGGACATCAGCGCGCGCAACGTGACAGCCACTGGTACGGTGAAAGCCGGCACT
GCTGACGTCCCGGCGAGACGTACACCGGAGGCTGGTTACGACCCGTGGTGACACGGGCTGGTACAACGAGAAATGGGGCGG
CGGCTGGTACATGAGCGACAGCACCTGGGTGCGCTCTGGATGAACAAGAACGTCTACACCGGCGGCGAGATGAAAGCGGGCA
AACTCACCGCCGAGGGCCGACGGAAGTCGGCGAGTACCTACAGCTCAAAGGCGTGGCCACCGAAGGAGCCAACTGCTCGCCG
AACGGGTGGCAGGCATCACAGCACCGGACTCTGGCTGTCTGCCAAAACGGGAAATGGGGACGAACGCCGCCTCCATGCG
CCTGAACACACCGCGCGCGTGTCAAGGACTGGTGTACGTTGCATGGTCAGGATAGCGCCATGGTGAACCTACGACTACGTC
GCTACGCGATCACCTGCGCGGCGGATTCTGCGCAGTGGGCTTCAACCGACATTTGGCACCAACTACTCGTTGCGGCTAATC
ACTGAGATCGGCCAGGCTTCAACTACCGGAACCTACAAGACCCCGACTCGACCAACGTGACCGTTACCTGCGTGAACCTA
G

54/118
FIG. 30I

Protein sequence: (SEQ ID NO: 179)

MRSTRSSGFISIELMIALVVIAIATAGGISVLMSYLDGLDEQHAAQQQQQVAKAAEKYLKDNFSTVLASAGATAPAVITVPM
LNTRYLPAGFRDNIYGOQYQVLARKPAANQLETLIVTTGGQVASELSIRRIAQLMGATGGYISKNTNTSIAQGAAWQVALSNF
GSAPGAGHLATALFFQDGAIANEYLYRNAVPGHPELNRMTTLDMGNNIAAAGAITASGNITTSADISARNVTATGTVKAGT
ADVAGETYTGGWFRTRGDTGWYNEKWGGGWYMSDSTWVRWMMKNVYTGEMKAGKLTAEGRTEVGEYLQLKGVATEGANCS
P NGLAGITSTGLWLSQNGKWGRTAASMLRNTTAGVIKDWCTLHGQDSAMVNYDYVRYAITCGGRFCAVGFNQTFGTNYSFGLI
TEIGPGFNYPEPYKTPDSTNVTVCVN.

RL079

DNA sequence: (SEQ ID NO: 58)

GTGAGTGTGAACCCGATCATCCAGGCTCAGTTCGTGACCTCTACCTCGGTGAAGGCTTCGCCGACGTGAAAGGCCTGGCCGG
CGCCGGCGCGCGCCGAGTCGAAGTGCCTCGCGAGTGGGAGTCGCACGTCCAGGAACCTGCTCCAGATCTGCAGGCAAACGCTGG
AGGAGCTGCAGGACCCCTGAGTTCGCCATCGTCGTGACGGCGTTCTGCTTCGCGTCACCCCTCCTCGAAGACGCTTTTCACTGGC
AGCGTCTTCGTGCTGCCCGGTCGAGCGCCCAATTGCGGGAGTTCCAAGAGATCGGCTATCCGAGCGAAGTGGTTTCCGCACT
GATGGATCCGAGTTGCAGGGCCTGGTCTGTTCTGCGGCGAGATGGCGACAGGCAAGACCAGCTCCGCCGCTCTCTGCTCC
TGGCCCGCCTGCAGGAGTTGGGCGGGTGGGCTGCGCCGTGAGGACCCGCGAGGAAACCAACCTCAGCGGTGAGCATGGGCTC
GGCCGCTGCATCCAGGTCAGAACCTCACGCGCTCAGGCGGATACAGCGAGGCCCTGCTGCGCACGCTGCGGGCGCGCCGA
CCTGGTGTGATTGGCGAGATCCGCGACGAGGACACCGCTACCAGGCTGCAAGGCCTCTCTGACCGGCAGCCTGGTGATCG
CCACCATTACGCGAAAAGCTGTTCATCAGCGCATCGAGCGCTTGGTGACGCTCGCCCGAGCCACTGGCGAGAAACGCCCTACGAC
GTGGTTGCCGAAGGCATCCAAGCTGTGATCTGCCAAGCGCTGGAGAGCGATGGTTCTCGCGCCGCTGACCGCCGAGCCACT
GCTGTTCACTGGCGACGAGCGCCGTCATGCGCGACAAGATCCGCCGAAAGGAGGCTCATCTGCTGAGGACGACCAAGCTC
GCCAGTCCCGGCAAAGCCTATGGAGATAA

Protein sequence: (SEQ ID NO: 180)

VSVNPIIQAFVDLYLGEFADVKGLAGAGARRVEVPREWESHVQELLQICRQTLEELQDPEFAIVVDGVLLRVTLLLEDAFSG
SVFVLRSSAQLREFQEIYGPSEVVSALMDPQLQGLVLCGEMATGKTSSAASLLARLQELGGVGCavedPQETNLSGQHGL
GRCIQVRTSRRSGGYSEALLRTLRAADLVLIGEIRDDETAYQACKASLTGSLVIATIHAKSCHQAIERLVTLAQPLARNAYD
VVAEGIQAVICQALESDGSSRRRLTAEPFLFTGDDGSPMRDKIRKEAHLLODDQARQSRQSLWR.

RL080

DNA sequence: (SEQ ID NO: 59)

ATGAGCACTACGCAACGCACCTTCCCGTCCGACGCAGGGCGGTTTCGTTTCCATCGAGATGATCATCGTGCTGATCATCATCGC
CATCGGGTCGGCCTGGCCCTGGCCGCAGCGGCTGGAATGTTAGTTCGTCCACGCCAACGAGGAACAACGCAACATCAGCG
TCATTGCGGCCAACGCACGCGCCCTGAAGACCTCTTCGGGCTACGGCTCCAGCGGTACCAACCTGATCCCCAGCCTGATCGCA
ATCAACGGCGTGCCGAAGAATGAGTGTCTCTCCGGCTGCTGTACACGCTCTACGGCGGATCGGTCACTGTCTCGTCCAC
CGCATGGGCTTCTCGATCACCACGCAAGTTGCCCGAGGACGCTGTATCAGCTGGCCACCAAGATCGCGAAGAACACCT
TCGAGCAGACCAAATCAACAGCGGATCCTCGATCACCAGGAGAAGTGACCACCGCAGCCGCGACCCAGGCCTGCAGCAGCGAC
AGCAACAGCATTACCTGGACCTATAGTTCGTGA

Protein sequence: (SEQ ID NO: 181)

MSTTQRTSRPTQGGFVSIEMIIVLIIIAIGVGLGLAAAAGMFSSSNANEEQRNISVIAANARALKTSSGYGSSGTNLIPLSLIA
INGVPMNSVSSGVVNVYGGSVTVSSTGMGFSITTSKLPQDACITLATKIAKNTFEQTKINSGSSITGEVTTAAATQACSSD
SNSITWYTS

RL081

DNA sequence: (SEQ ID NO: 60)

ATGGGGGGCTTCTGGGAGCAGTTGCAGTTCGCCTTCTACAGCAAGCAGTTTCGGCCGCAAGGAACGCCTGCGATTCTACGAAAG
CATGTCCACCCTGCTCGAAAACGGGGTCCCGTTGAAGGATGCTGTGGCAGAGGTGCACAAGATCTTCGCTCATGAGGGGAGC
ATCCGTTTTCATCCGGTGCCCATCGCCAGTCGCGAAGCGCTGATGGGGCTGTCCAACGGCAAGCGTCTGGCCACCGCATGGCG
CTCTATCTCCCCGCCAGGAGCGAGCGTTGATCGAGGCCGCGGAGATGAGCGGCAACCTGGTTTCAGGCATGGGCGATGCCGT
CTCCCTGGTTCAGGGCCAGGCCAGGATCCGCGCCACCATCTGGCAGGCGCTGCTCTACCCCTCGGCGCTGTCCGCCATGATGG
TGTTCTGCTGTGCATCGTGGCCTATCGCATGGTCCCCAGCCTGGCCAGGCTCTCCGACCCAGTCACCTGGACCGGCCCGCTC
GCCACGCTCAACGCCATTGCCAGCTTCGTACAGGACCTGGTATCTACGTTCTGGTTCGCGCTCATCACCCCTCACGGTGGTGGT
CATCGTCACGTTGCCGACCTACCGCTGGAAGGCCGGGTCTGGCTGGACCGGACGCTGCGGCCCTGGTCCATCTACCGCATGC
TCCAGGGCACCACCTTCTGCTGAACATGGCGGTGATGCTCAACGCCGCGATACGCCCTACGACAGCCTGGCCAGCATGATC
AAGATCTCCCCGCCCTGGCTGAAGCAGCGCTTGAAGCTGCCCGCTACGGCGTGGGCTGGGCCAGAACTTGGGTGTTGCCCT
TCGACGCGCGGTCACGATTTCCCGACCGGCGAGGCCATCAGTACCTGTGCATCCTCGCCAACCGGGGAGGCTTCTCCGAGG
CGCTGGTCAAGTTCAGCCCGCGCTGGCAGGAGACCAGCCTCAAGCAGATCGAGCTGGCCGCGGGGCTGGTGAAGAAGTTCGCC
CTGATCTTCATCGGCGCGCTGATGATCTGGTCTGCTCGGCGCTACCAGGCACAGCAGCTCATCCAATCCATGAACCACTG
A

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND
PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September 12, 2003 Serial No.: Not Yet Assigned

Page 54 of 118

Customer No.: 21559

55/118
FIG. 30J

Protein sequence: (SEQ ID NO: 182)

MGGFWEQLQFAFYSKQFGRKERLQFYESMSTLLENGVPLKDAVAEVHKIFAHEGQHPFHPVAIASREALMGLSNGKRLATAMALYLPAQERALIEAGEMSGNLVQAMGDAVSLVEAQARIRATIWQALLYPSALSAMMVFLLCIVAYRMVPSLARLSDPVTWTGPLATLNAIASFVTGPGIYVLVAVITLTVVIVITLPTYRWKGRVWLDRTLPPWSIYRMLQGTTFLLNMAVMLNAGIRPYDSLASMIKISPPWLKQRLAARYGVGLGQNLGVALRSAGHDFPDRQAIQYLCILANRGGFSEALVKFSRRWQETSLKQIELAAGLVKNFALIFIGALMILVLLGAYQAQQLIQSMNH.

RL082

DNA sequence: (SEQ ID NO: 61)

ATGACGAACCTTCAGATTGCCGCGCTTGCGCAGCCCTCCATGGTGACCCAACCTGCTCACCGCCGACGGTGGTGAATGGGAGGTATCGAAGCACCTGCAGGAAATCATGGCTCTGGCTGCCGACGGCAGCTCTATCTATCGGAGAGCCACCAGAACGACATACAGTTCTGTCTGTTTCATCGACCGTCTCGATCGCCGTGGCTCCGATACCAGCTCAACCTCACCGACCTGCAGACCATTACCAGCTTTACCGCGCCGTGCGCATGGACGGCCTGGTTCGATAGCGATGGCCAGCGCGCCACCCAGATGCAGGAGCGCGTGGTCAAGATCATTCGTAAGGCCACTGAGCTGCGCGCCAGTGACGTGCATTTTCGTCTGAGTCCCGCCGGCACCGGCAGCAAGATCCGTTTCCGCGTCGACGGCCTGCTGAAGACCGTGCAGCAGTTCCGCGAGCCAGGAGCTGCACGAACTCTGTGCAACCATCTACCAATCCATGTGCAGCTGGCCGAGCCACTGTTCAAGCCGCAACTGGACCAGGACGCGCGGATGAGCCAGACCTTCGTGAGAAGCTCAACCTGTTTACGTGCCCGGATCGCCACCCCGCGCTGCCGGGGGTTCTGATGATCCTGCGACTGCTCTACGACGACACCGGCCTCGACAGCCTGGAGCAGCTCGGCTACCTGCCCGAGCAGAACGCACCTGTTTCGATCGCATGATGCGTATGCCCTACGGCATCAACATCCTGTCCGGCCCCACCGGGTCAGGAAAGTCGATGACCTTGAAGGTCACCTGGAAGGCCTCGACAAGCTCCATGGCGGATCCAAGCATATCTGACCATCGAGGATCCGCGCGAATACCGCATTTCGCGCGCAAGGCATCAACCAGACCCCACTGGTCTACGACGCCACCGACCCAGACGCAGAACGCCAGGCCTGGGCCGCGGGCATCGCCAACGGCATGCGCCTGGATCCGGACTACATGATGATCGGCGAAGTACCGGACCTCTTCGCGCTGTTCGCGCGCTTCGCTGGTGACCGGGCACGGCTATGGTCGACCCTGCACACCAACAGCGCGATCGGCATTGTCCAGCGCCTGAAGGACCTGGGCGTCGACCCCGCTTGTCTGTCGATCCGGCCTGCTGACCGGCTGATCAACCAGAGCCTGCTGCCAAGCTCTGCCCCACTGCAAAGTGCGCTTCCAAGACCACCAAGACCAACTCGCGCCCCGACTGTGTCGAACGGGTCCGACGCTTGACCGATGTTTCCAGGTTACCTCAAGGGGCGCTGCTGCCAGGCCTGCCGTGGCTCCGGGGTCAACGGCCGCTCGATCGTCGCGAGGTGGTTCTGCCCCACCTCGCCTTCATGCGTGTGTTGCGCAAAGGCGGGCCAGCCGAGGCACGCAACTACTGGGTCAAGACCATGCAGGCATCACCAAGCACGCCACGCCATCCGCCGCATCAACGAGGGCATGTTTCGACCCGAGATGGTCGAGGATTTTATTGGGCCACTCGACTTCGATGAGCATCTGCTCGACGACAGCTTCTACTCGCAGGAGGCGTGCTGA:

Protein sequence: (SEQ ID NO: 183)

MTNLQIAALAPSMVTQLLTADGGEWEVSKHLQEIIMALAADGTYLSESHQNDIHVLSFIDRLDRRGFRYQLNLTDLQTIHQLYRAVAMDGLVDSDGQRATQMQRVVKIIRKATELRASDVHFVVPAGTGSKIRFRVDGLLKTVEQFRSQELHELCAATIIYQSMCDVAEPLFKPQLDQDARMSQTFVEKLNLFSAIRIATRPRAGGFLMILRLLYDDTGLDSLEQLGYLPEQNALFDRMMRMPYGINILSGPTGSGKSMTLKVTLGLELDKLHGGSKHILTIEDPPEYRIRGEGINQTPLVYDATDPDAERQAWAAGIANGMRLDPDYMMIGEVERDLFAAFAFRGAMTGHGLWSTLHTNSAIGIVQRLKDLGVDPLLLFDPALLTGLINQSLLPKLCPHCKVRFQDHQDQLAPDLVERVRLTQDVSVQVHVKGPGCQACRSGVNGRSIVAIEVVLPTLAFMRVFAKGGPAEARNYVVKTMQGITKHAHAIRRIINEGMFDPQMVEDFIGPLDFDEHLLDDSFYSQEAC.

RL083

DNA sequence: (SEQ ID NO: 62)

ATGCGAACTGAGCCGATCGGCATGGCCGTGGCCGTGCTCTTCTCCTCGCGTCTGGCCAGGCCTGCGCTGGCACCGTTGGCGAATCTGCGGAGATCCAGGCCCCAGGCCATCCTCACCGAGGCCAAGGTGCGCCTGGCCACGGCGCAGCGGCAATTGGAAGGCAAAGGCGAAACCCGGCCAGGTTCGTTCAGCGCCCCAGGGCAGAGCTTCGCCATGCGCGTGGCCGGCGCGCGCCGACGATCACGCAGCCGTTCCGCGCAGTGGTGCAGGACCATCTACGGCGCGCGGAGGATGCTGCGGACGTTCTTGTTCGCGGGCGGGTACGAGGTTGACGCGCCAGCGCGCGGAGCTGCCTGGCAAATACCGCGTCGAGTCAATCTCGCTGGACCGAGTCTGCTCACCGACAAGGACGGCAACCGCGTGGCCGTGGGCTTCTCCAGCGTTGCCCCACCCAAGCCTCCTCTACGGCCCCAAGGCGCCTCGGTTCCGCGCGCGCTGACCGCAGCCGTTTATTCAGTAG

Protein sequence: (SEQ ID NO: 184)

MRTEPIGMAVAVLFLASGQACAGTVGELAEIQAQAILTEAKVRLATAQRQLEGKGETGVVSAQGGTFAMPVPAAPPTITQPVPPVVRTIYGAGGKMTATFLFPGGYEVDAASGAELPGKYRVESISLDQVVLTDKGNRVVPVGFSSVAPTQASSTAQGASVPPALPGAVPQPFQ.

56/118
FIG. 30K

RL084

DNA sequence: (SEQ ID NO: 63)

ATGGAGAAGCCTGACCTCGGCAGCCGTGGACCAGACGTCTCGATCCTGAGCTACCACGGCAACAAGTTCGTGACGGCCCTGTT
CTGGCGGCCGCTGTCCAGCCAGCGGCAGTACATGAAGGAAGCGCGCAAGCTGGGCAAGGAAGAGCATCTGGACATCGTTGCCA
TCCGCCATTACCGACGGTGATCCAGGCCGGCTTCGTTTCGAAGTCGCAAGGCGCAGTCAAGGGGATGTACTCCCTGGCCTCG
GCGCTTTCAGGCCAGTTCGACGGCGACTTCTGGCCTGCTGGAAAGTCGACGAGGACCGCTACGCGCTGGTCCGACGCTCGA
TGGCGCGATTGTCCCGGGCAGGATCTGGTCACCACCCTGACGAGGCCCGGACCGGGTCAGGAAGCTCTCTACGCGCGCG
TGCTGCGAAACGCACAGGTCTTCGTTCCCGAAGGGTTCGATTTCCTCCGTCGACATCGAGGAACTGCTCGCGCGC
AAGCGCTGCGGCGCGACTACCGCTCCGGCAACTCACCTTCGGCTTGTCCGCCAGGGAGTGGACGGCAGTGGCCCTGCTCGG
TTGCGTGGTAGGTGGGTGCTAACCCTACTACCTATGGAATGCCCCAGGAGAGCTCGCCAGGCAAGCCGCGCTCCTCG
AGGAGCAGAGGCGCTCGCCGAGCTGGCCGAGAAGAAGCCAGGCAAGCAGCCGCTGGACCTGGCGTCATTGCAGAAGCCT
TGGAGCTCATACCTGACCTCGAGGACATGCTACGCGCTGTAGCAAGGCAACGGGGTACTGTGCTGTGATCCAGGGCTG
GCTCTTCGAATCCAGCAAGTGCACGGCAGGGTCTGGTGCACCTACCACCTACCGGCAACAGCAGCAGCCGACCTGA
CAGCGGCCAGCCAGCACCTGTTCCGCGACCGCCCCGCTTCGTCATCGACAACGGCAACACCGCGGCCCTGAAGGTCGATCTG
AAGGTGGCCATCGGCAGTGATGAGCCGCTACTGCCGGCGGACGACGTTCTGCAGGCGCTGACGAGCCACCTGTACCGTCAAGG
GGTCGAGCCCAAGCTGTGATCAGCCAGGAGCAACTCCGCCCTCCTGGCGCGGAAGCTGCGACTGAACAGCAAGTGGTGT
TGCTTTCCTGGAAGAAATTCACCTTCAGCGCCGAGCCCGCTCCCGGCGAGACCTGACCTTCAGGGGCTGCCCGCTGCCGT
GTCCGCATCACCAACCTCGAAACCAGCTCAAGGACAGCCAGTTGGACTGGACTGTACAGGAGAAATCTATGCGAACTGA

Protein sequence: (SEQ ID NO: 185)

MEKPDLGSRGPDVSI SYHGNKFVSGLFWRPLSSQRQYMKEARLKGEEHLDIVAIRHSPTVIQAGFVSKSQGAVKGMYSLAS
ALSGQFDGDFLACWKVDEDRYALVATLDGAI VPGQDLVTTLDEARDVRKLS TRGVLRNAQVFPVPEGDFDPVKDFDIEELLAP
KRLRRDYRLRQLTFGLSAREWTAVALLCVVGSLTAYYLWNAHQEELARQAALLEEQRRLAELAEKNAQAKQPLDLASLQKP
WTLIPDLEDMLRACSKATGVL SLSIQWLFESSKCDGRVLVATYHRTGNSTAADLTAASQHLFADRP AFVIDNGNTAALKVDDL
KVAIGSDEPLLPADVDVLQALTSHLYRQGVPEPKLSISQETTPPLPGAEEATEQQVVLPSWKKFTFSAQTRL PADLTFQGLPAAG
VRITNLETTLKDSQLDWTVTG E IYAN.

RL085

DNA sequence: (SEQ ID NO: 64)

ATCGTGTGCGAAGCTACGGCAGATTCCGCGTCTACGATCGCAGCGCAGGTGCGCAACACCCGACCGGATCGGCGCGATACGGT
GGTGTCTCCGACAAACCTGGGTGAGCAGCAAAACCCCTAAGCGTTTCGCACACCTTGTCCAGTGACTGCATCGTGACGTGGC
GCCCTGCAGGCGCAGCGTGCCTGCAGGAGGCCGCCAGGAAGTCATCAACCAATGCCACATGGCGGTGAGTATCACGCCCGAC
GCGTGAACCCGGCCGCTTCGCGGTGCAACCTCAGCAGCGCGCGAGCAACGCCCGCCGCCATCCAAGGCGGCCAGGACAT
GGCCACCATTGCTGTTTCTGCTCCGTCGCAACGGCATGTGCTCGGTGCCGCGCGCAGCATGGGGTCGAGCTTCGGGTCTCT
ACGGTCCGCGTCTCTTACAAACATCAATGGAACGGCAAGTCAGCGGGTTCCTCGATCTCATCGCCGCGCCAGCCGCGGTG
TCCTGGCGCTACAACCAACCGAGAAAAGGGTCGAGTTCCTACTACCTGGACACTCGGACCTTCGCGATGTACGCTTCGACGA
CGTCAACACGGTGGACTCCACCGTGCCTTCGGTATGACGACGCGCCGCGGCATCAGCGGGGACGGTCCGGATCCACCGGAC
AGAAATGGCAGCTCCGCGATCAGCGGCGACTCCGCGCAGCAAGCAGACCAACAGCTCGGAGCTGAAGACATCGATCCTCAGCGAC
ATCGAGAACAGCATCAACTCGATGCTGACGCGCGAGCATGGGACGCATGTGCTGTGCGGTGCCACGGGCAACCTGACCGTCA
CGACCGTCCAGAAGTCTCAACCGTGTCCAGCAGTTGGTCAACCGAGAGAACGAGAGCATCACCAAGCAGGTGCTGCTGAACG
TCAACGTGCTCTCGGTGCGCCTGACCGACAAGGATCAACTGGGGATCGACTGGAACCTGGTCTACAAGTCGCTCAACAACAAG
TGGGGCATCGGCCTGAAGAACACCATGCCGGGCATCGATCAAAGCGCGATCTCCGGCTCCGTGAGCATCCTGGATACCGCAA
CAGCGCCTGGGCAGGATCCAAGGCCATGGTCCAGGCGCTGGCCAGCAGGGCCGCGTCTCGACCGTCCGATCCCCGTCCGTGA
CCACGCTCAACCTCCAGTCGGCGCGCATCCAGATCGGCGCTACGACAGCTACCTGGCCTCCAGCCAGATCTCAACGTGCGC
CAGGTCCGCGCAGTACCACTCGCTGATCCCGGCGCGCTGACACGCGCTACAACATGAGCCTGCTGCGCTTCGTGATGGAAG
CGGCGAGATGCTGCTGAAGATCAACATCAACATGACCTCCCGGCGGACGTTTCGAAATGCAGACCAGCGGGGACTCCAAAGCCC
AGTTCGCGAGCTACGACATACAACCTGTTGACGAGAAAGTACGCTGTCGCGAGCGCGAGACCTTGGTACTCTCCGGCTTCGAC
CAGACCACCGAGGACACCAACAAGGTCCGACCGGCGACGCTGGGTCTCTCGGTCTTGGCGGCGGGCTGACCCGCAATACCAA
GCGCGAGGTATCGTGGTGTGATCACCCCGCTCGTGTGGGCTGA

Protein sequence: (SEQ ID NO: 186)

IVCEATADSASTIAAQVRNTRPDRRD TVVFS DKPWVSTKPLSVSHLSSDCIVTWRPAGAASLQEAQEVINQCHMAVSITPD
ALNPAAFAVQPQQRASNAPPPIQGGQDMATMLFPASVANGMSL GAGGSMGSSFGSYGPRSLYNIKWNGKVS GFLDLIARAGV
SWRYPNTEKRV E FYLDTRTFRMYAFDDVNTVDSTVRSGMTT AGISGDGSGSTGQNGSSGISGDSGSKQTTSSELKTSILSD
IENSINSLTPSMGRSLSRATGTLTVDRPEVLNVRVQQLVNRENESITKQVLLNVNLSVALTDKQLGIDWNLVYKSLNNK
WGIGLKNTMPGIDQSAISGSVSILDTANS AWAGSKAMVQALAQGRVSTVRS P SVTTLNLQSAPIQIGRYDSYLASSQISNVA
QVGS TTS LIPGAVTSGYNMSLLP FVMESGEMLLKININMTSRPTFEMQTS GDSKAQFP SYDIQLFDQKVR LRSGETL VLSGFD
QT TEDTNKVG TGDAGFFGLGGGLTRNTKREVIVVLITPVVLG.

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND
PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September 12, 2003 Serial No.: Not Yet Assigned

Page 56 of 118

Customer No.: 21559

57/118
FIG. 30L

RL086

DNA sequence: (SEQ ID NO: 65)

ATGACCAGGCAGTTGACCACTCTCACGCTGTGCTGCTCGCCAGCTGCACGACCCACAAGGCTGAGCCGGCCAGGCCAGC
CTTCGACAGCAGCCGAATCCAGACCTGCTTTCTCCGGACCTGTATCCAAACGGTGTGCAGCCGGAGAAAGAGCCCGTAGTGC
GCTATGGGCGCTACACCTGGTCAACCGCAGCCTGATGCCGGTCAACGCGACCTGATGGCCAGATCATCGACGTAACCATC
CCGTCGAGCATGAACCCGAGCGTCAAGGACGCCATGCAGTACGTGATGAGCCGCTCGGGTTACTCGCTGTGCCCGGCAGACGC
CGGTATGTGAACATCTCTACACCCGGCCGCTGCCGCGAGTCAAGTACAAGCTCGGCCGATGACCCCTGCGCAACACCTCC
AGGTCTCTCCGGCCAGCCTGGCAGGTTAAGGTGACGAGTTCGCGCGGAGGTCTGCTTCGTGCTGCGCCCGGGCTATCAA
CTTCCCCCGGCGCGAGGCCGAAACCGTCCAGCAACTGTATGCGAAGCCCGCTGCCCAACTCCGCGCGGCTAGCGCAACC
CTCTCCACGGAGAAAGTCAACGCTGGAGTTCGCCCCTCGGTGCGCTCGGTGCCGACACCGCGCGCCGATCACAACCAGCC
ACGCTCCGGCCAAGAAGCTGAATCCACCACTGTGCTCCCCCAGCCGACCGGCCAAGGATGGCCACCCCTCTTCTCTCTCC
GCGGTTTCGGCACCGACCAAGCCTGCGGCTCCGCGGTGAAGTCCACGCGCCACTCCACCCACCGTGGCTTCGCGCCACCC
GGTCAAGGTGCTCACGCGCGCGAACCAGCGCGCGCTGGCACAGGCCTGGTCAAGCGAGACGGGATCAACCTTGCAGCGACA
CCTTGAAGCTTGGGCAAGCGCGCACGCTGGACCGTCCGCTGGGAGCCGAGGATCTCAACTATCCGATCGAGGCTCCACTG
ACCTTCCACGGCTCCTTCGAGGACGCGGTATCCGAGCTGTTCCTCTGTATGACGCTGCCGAACGGCCCTTCTGGTGAACGC
CAGCCGCGCGCAGTCCCTGATCATCATCAAGGAGCGCAAGAACTGA

Protein sequence: (SEQ ID NO: 187)

MTRQLTTLTLCLLLASCTTHKAEPARPAFDSSRNPDLLSPDLYPNGVQPEKEPVVRYGRYTLVSTQPDAGQRDLMQIIDVTI
PSSMNPVSKDAMQYVMSRSGYSLCPADAGHVNIYLTRPLPAAQYKLGPMTLRNTLQVLSPAWQVKVDEVARQVCFVLRPGYQ
LPPAPRPKPVQQLYAKPAAPTTPPAVAQPSSTEKVVSTLESPIVVASVPTPAPITTSHPAKKPESTTVLPAPAKDGHPSPP
AASAPTKPAASAVKSTPPTPPTVASAPPVKVLTPEPSRPLAQAWSAETGSTLRDLTLEAWAKRARWTVRWEQDLNYPLEAPL
TFHGSFEDAVSELPFLYDAAERPFLVNASRPQSLIIKERKN.

RL087

DNA sequence: (SEQ ID NO: 66)

TTGAGCTTTAAATACTATTGGGCTAAATTTTCTGGGGAGCTTTCTTCTTTGTTTGTAGTCGCTTGAAAGGCTCCGTATTTCC
AAGCCTGGCATCAGTTAACCCCTTGGTAGTGGCTGGATTCACTACTATCCTGTTTCTTCTCGGTAAGGCTTGTGAAGACT
TCGCTTTAAATATACGAAAAAGAGTTCTGGGTACAGGTTTCTTCTCCGAAACCCCTGCAAAACAGGATTGTATCGAGTC
TTTATTTGGCTTGTATTGTTTCAATTCCCTTGGGGATGATTTTTTTATTCTATAAATACGGAAAGGCCTCGTAG

Protein sequence: (SEQ ID NO: 188)

LSFKYYWAKFFWGAFFVLVAWKGSVFPSLASVNPLVVAGFSTILFPFSVRLVEDFALKYTEKEFWVTGFFSETPAKTGL
YAVFYLACYLFSIPLGMIPLFYKYGKAS.

RL088

DNA sequence: (SEQ ID NO: 67)

ATGTCCAATGACAACGAAGTACCTGGTTCCATGGTTATTGTGCGACAAGGTCCAGACGATCAATACGCATACGAGGTTCCCCC
TATCGATAGCGCGGCCGTTGCCGGGAATATGTTTGGCGACTTGATTCAAAGAGACATATATCTACAGAAAAACATTTATTATC
CAGTCCGATCCATTGTTGAACAAGGAACAAAAGAAAAGAAGGAGATCAACAAGAAAGTATCTGATCAAGTCGATGGCTTGCTA
AAGCAGATCACTCAAGGAAAAAGGGAGGCCACAAGGCAAGAGCGAGTTCGATGTCATGTCGGCAGTCTGCACAAGATGGAATC
TGATCTTGAAGGATACAAAAAGACCTTTACCAAAGGCCCATTCATTGACTACGAAAAGCAGTCAAGCCTCTCCATCTATGAGG
CCTGGGTCAAGATCTGGGAGAAGAATCTTGGGAAGAAAGAAAGTACCCCTTTTCAGCAGCTTGTAGAGATGAACTGGAG
CGGGCGGTTGCCTACTACAAACAAGATTCACTCTCTGAAGCGGTAAAAGTGTCAAGACAGGAGCTCAACAAGCAAAAAGCGCT
AAAGGAAAAAGAGGACCTCTCTCAACTGGAGCGGGACTACAAAACCAGAAAGGCGAATCTCGAGATGAAAGTACAATCCGAGC
TTGATCAAGCGGGAAGTGCTTTGCTCCATTGGTCAGTCCAACGCCAGAGCAATGGCTTGAACGTGCCACAAGACTGGTTACG
CAAGCAATTGCTGATAAAAAGCAGCTGCAGACCACAAACAATACTCTTATCAAGAATGCCCAACCCCTCTAGAAAAGCAGAA
AGCCATCTACAATGGTGAGCTACTTGTGGATGAGATAGCCAGTCTACAGACCCGCTTAGATAAGCTGAACGCCGAAACGACAC
GACGCAGGACAGAAGCAGAACGCAAGGCGGCCGAGGAACAAGCGTTGCAAGATGCTGTTAAATTTACTGCCGACTTTTATAAG
GAAGTAACTGAGAAATTTGGCGCACGAACATCAGAGATGGCGCACCAACTGGCCGAAGGCGCCAGGGGGAAAAATATCAGGAG
TTCCGGCGGAAGCAATCAATTGTTTGA AAAACAAGGATGCGTTAAATAAAAAAATAGCCTTAAAGATAGGCAAGCCATTG
CCAAAGCCTTTGATTCTCTAGACAAGCAGATGATGGCGAAGAGCCTTGAGAAATTTAGCAAAGGCTTTGGAGTTGTAGGCAAA
GCTATTGACGCCGCCAGCCTGTACCAAGAGTTCAAGATATCTACGGAACCGGGGACTGGAACCAATTCTTTGTAAAAGTTGA
AACACTAGCTGCTGGTGGCGGCCAGTTGGCTTGTGGGTATGCAATTGCCACGGCAACGGCCACTCTATAGGCATCCTGG
GGTTCGCACTGGTAATGGCAGTTACCGGGCGGATGATTGACGAAGGCCTTCTAGAAAAAGCAACAACCTTGTAATGTCCATT
TAA

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND
PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September 12, 2003 Serial No.: Not Yet Assigned

Page 57 of 118

Customer No.: 21559

58/118
FIG. 30M

Protein sequence: (SEQ ID NO:189)

MSNDNEVPGSMVIVAQGPDDQYAYEVPPIDSAAVAGNMFGLIQRDIYLQKNIYYPVRSIVEQGTKEKKEINKKVSQVDGLL
KQITQGKREATRQERVDVMSAVLHKMESDLEGYKKTFTKGFPIDYEQSSLSIYEAWVKIWEKNSWEERKKYPFQQLVRDELE
RAVAYYKQDSLSEAVKVLRLQELNKQKALKEKEDLSQLERDYKTRKANLEMKVQSELDQAGSALPPLVSPTEQWLERATRLVT
QAIADKKQLQTTNNTLIKNAPTPLEKQKAIYNGELLVDEIASLQTRLDKLNAETRRRTEAERKAAEEQALQDAVKFTADFYK
EVTEKFGARTSEMAHQLAEGARGKNIRSSAEAINSFEKHKDALNKKLSLKDRQAIKAFDSDLQMMAKSLEKFSKGFVVGK
AIDAASLYQEFKISTETGDWKPFFVKVETLAAGAAASWLVGIAFATATATPIGILGFALVMAVTGAMIDEGLLEKANNLVMSI
.

RL089

DNA sequence: SEQ ID NO: 68

ATGAACCGTCCACGCTGGTTAATCGTACCTCCGCGACACCTTCGACGCTTCTGCAGCGGGCTATCTTCGACGGCTACGACTT
CGGCTTGAAGATCCCCACATCGCAGGCAGCAATCGCGCGTGTCTGGAGCTGTCCGGCTTCTTCATCAGCGCCCGGGAGCATC
CGTTGCACCGCTACTGGCGGGTCCCCAAAGGCAAGCTGCTGCCTGAAGTGGACACTCTGTACAACCGTCTCGCCGAGCTAGCT
GGAGGCCTTCACTCCCAGTCTCTGGCGGGAGTTCAGCTCCTTGGTCTGAATCCGCGCAGGCCTCGCTTGACCGACAGGCCTTCAC
CTGGGGGATGCTGCTGCGCATCGCGCCCCCTGGCCGAGGGCGGCGTCTACTGTCTAGGCGAGTTCATCCTGGTGTGTGGCGG
TGGCTCGGCGGATGCGCGGGGTATTCTGCGCCCATCGAGTTCTTGGCGCATCGACACCACTCCCGAGCTGCTCCGAAGCAAC
CTGATTCTGGAGCTTGGCTCGCCGAGGAACAATTGAGATTCTGGATACTGTCCAGGAGCTGCTCAGCGACGGCAGCTTCGC
GCCGTCGACCGAGCTGCCAGCATGAGCATCGCGGTCCACAGCAGGAACCGGCAGCGCCATCCCTGGAGGACGAGTCAGCCT
CTGACATCTACCTCGCCGCGGTGCCGAGATCGAGCGCACCGAGTACAGCTCGGCTGATATCGAGGCGGCGCTTCAGGGCTAC
TCTCTACTGGCCACAGCCTGACGGCATCGCTCATCTGCTGCAGAGAACCAGCGCCTTATTGGCCGACGACATGGGATTGGG
CAAGACCGCCAGGCGGTTCATCGCCGCTTCGATCCGCGCGGCGGCGAGACCAATCCTGGTTCATCACCTGGCTACCTGCTGA
TCAATTGGCAGCGGGAGATCCAGGAGGTCTATCCCTCGGCCACCGTGGCCATCCAGCAGGACACCCAGAGGCGCAGTGGATC
CTCGTCAACTACGAGCAGTTGAGCCCTTCGTCGCCAACGCTTCGCGCTTCGCCGTGATGGTTCATCGACGAGGCGCAGCGGAT
GAAGGAACCGACGGCGCAATGCACGCGGCACGGTTTCGACATTGCCGCCCAAGTGCCGAACCGCTACCTGCTTACCGGCACGC
CGGTGCTCAACCGCGAGACAGAGCTGCACACCCTGCTGCGCCTCTCAGGCCACCCCATCGGCCAACTGCCGCTGAAAGAGTTC
TGCGACCGTTTCGCGCGCAACCCGGAGTTCCGCCAGAGTCTGCGGGCGGAGCTGGGTGACTGGATGCTGCGCAGGCGCAAAGA
TGTGCTGCCAGCCTCAAGGGCAAGCAGCGGCAGTTGCTGAAGGTGGCCCTCTCCACCAGGAACGCCAGCAATACGACGTGC
TGCGCCTCGAGGACCGACCGGTCTTCGCGCGACTCGGCGCGCTGCGGCGTTACCTGGAAACGGTGAAAGTTGCGGTGGCGATG
GACCTGTTGAGCGAGCTCGACGCAGAGGACAAGGTGATCCTGTTCTGCGAGTTCAAGCCGACCGTGGCTGCGCTGAAGGAACT
CTGCGAGCAGGCGGACACGGCTGCGTCACGCTGGTGGGCAATGACTCGCTCACCAAGCGGCAGAAAGCGGATAGATCGCTTCC
AGCAGGATCCCGACTGCCGAGTGTTTCATCTGCACTACGGCGCGCGCAGGACGGGCAACAACCTCACTGCGGCGAACTACGTG
TTTTCTCTCGGCCTGCCCTGGACTCCCGGTACGACAGGAACAAGCCGAAGACCGCGGTACCGAAACGGCCAGCTCCGCATGGT
CGTGGTGAAAATCCCACTGGTTCGAGGCCACGATCGACGAGCAACTGTGGCAACTGCTCAACGCGAAACGCCAGGTTGCCCAGG
ACCTCATCGAGCCCGAGCAGGTTCGACGGAACCGCGCGCTTTTAGCCGCAAGCCTAACTGGATAA

Protein sequence: (SEQ ID NO: 190)

MNRPLVNRTSATPSTLLQRAIFDGYDFGLKIPYIAGSNRALLELSGFFISAREHPLHRYWVRPKGKLLPELDTLYNRLAELA
GGLHSQSWREFSSLVESQAQLDRQFTWGMLLRIAPLAEGGVLLSGEFHPGVVAVARRMRGVFLRPSSSWRIDTTPPELLRSN
LILELGLAEEQFEILDVQELSDSGSFAPSTELPMSIGGPQQEPAAPSLADESASDIYLAAPVEIERTEYSSADIEAALQGY
SLLAHQPDGIAHLLQRTSALLADDMGLGKTRQAVIAASIRAAGRPIVLITLATLLINWQREIQEVYPSATVAIQDTPAQWI
LVNYEQLSPFVANASRFVAVMIDEAQRMKEPTAQCTRHGFDIAAQVFNRYLLTGPVNLNRETELHTLLRLSGHPIGQLPLKEF
CDRFAGNPEFRQSLRAELGDWMLRRRKDVLPVSLKKGKQRLKVALSTEERQQYDVLRLLEDPRPVFARLGLRRLRYLETVKVRVAM
DLLSELDAEDKVILFCEFKPTVAALKELCEQAGHGCVTLVGNDSLTKRQKAIDRFQDDPDCRVFICTTAAAGTGNNLTAANYV
FFLGLPWTGQEQEAEDRAYRNGQLRMVVVKIPLVEATIDEQLWQLLNKRQVAQDLIEPEQVDGNRALLAASLTG.

59/118
FIG. 30N

RL090

DNA sequence: (SEQ ID NO: 69)

GTGGCACCTCTCGACAACGCCCCCCTAGCGGGCGCTACAGGATCCATCCCTGGCCCCGCTACAGCGAGCGGCAGCTCGCCGT
CGCCAACACCTGGGCAACACATTTCTCCCTCGCAGGGACAGCTCGAACCAAGTTCATCCGTCACTACCTGCGCAGCACATCTA
CGACCAGGTGCTGGTGCATCACAGTCGCTGCTGACAACGGTGTGCGTTACACCATCATGCGTGCAGGGCCGCTACTCCAGGTA
TTCGACGGTCAACTAATTGGTGGTGGGAGTGCAAGCCTGCCATCGTATCCCGCAAGCACGCCGTCTCGAGCAGGGGCCCTT
GAAGCTGCTACAGCGCCTTCAAAAGTTCGACGACGAGTTCGTGTACTCAGCTCATACACAAAGCGAGCGCACGACCTAGCCA
CACAGATGGCCAGGGACGATCTCGGACTTCAACATCGCCTCGTGTATCCGAGCCACAGCAACAAGCGCTACTACGCGCCAAGG
CACCAGTTCTACTTGAAGCAGATCGGAGCGGTCTTGCGAACCTTCAGACAGGTCCTGGACCAAGACCTGCTGTTGCCATCCG
CTCGGTTCCGTGCCTCTCGCCCCAGCTCTACAACCTGGTGGTCAAGGCGACCAAGTGCGCCGGCTGCAAAATGCTGAAGGCTC
AGCCGGTCTTGACGCCGCTACTGGTGGATTGCGAGGAGGGAGTCTGGCCTCACACGACGACCAACGACAACGGCGAGAGCATC
CGCCATTACCTTCCCTTGGCCCTTTCCCCAGCTTGACAGTGAACGACCGCAGGCCGCCGCTATGCCATGCGACTTGTACCTCGA
TATGGGCCGTATTCTTGGGCAGGTGCGGACGAAGGAATTTCCGTTCATCAACTTTTTCGCTGGCTATTTTCAGGCGCCGCGGG
CCTCGATTGATTTCTTAGTCACGTCACTCCCGCCGTGCGGGAGGAGCTCTCTTCCATCGCAACGGGAAGGCCGACATTGCG
GGATGGCATGCTCTCTACTGGCGGCATCGCTAGGTAACCGCGCGCCGATCACTCGCGCTCAATGGACAGCATTCTATGCCGC
CTACAATGCGATCCCTTGGCAAGTTCACAACGCCAAGCCGACTACAACCGTCTCTTCAACGGCTGCCCGTCCGATTGGCAGG
ATCCGGCATGGCTTGCAATCACTGCACGGCTGAGAGACATCAAGGAGTTCTATACCGCCCTCGACCAAGGGGAACCTCACAGGTT
GTTCCGCGAGGCGCGCAGCGCCCTGAAAGCGTATCTGGGTCAATTGTACCTACCGACAAGCTGGCAACCTGGTGGACGACTACCA
CCAGGTCCAGAGGGAGTGGTGGCCGAGTGCAGAGCAGCCTGCCGATCTGGTGCACACCGACGAGTACACCACCTGGGAGG
GAATGCTGTCTGTGCGTCTTATCGATTGCCCTAATGGACTGCAGATCGTCGAGCTCCGCTGTCTGCGGACCTATATGCCGAA
CATATCGCTCTGGCACATTGCATCGATAGTACGACCAGGCCGCTACCGAGGAGACTGCCGACTGCTCTCAGTACGTGAGGC
TGTCGTCGCGCTGGCCTCTGCCGAATTGGAGCTCAGGCGTGAGCATGGCGAGCCTATAGGTAGGCCCTGGAGTCCCAAGCACC
TTTCCACGGTGAACCTGCGCGAATTCGATAATGCCCCGCTGCCGACCGACTCGCCTGCCGGCCAGGCATACCGCTGGTTCATG
GAACGAATTGCTCTGGAGCCATAGCGACGAACCTGAACTGGCCCGACATGACCGTCCACATGACGCGCTTCGCCAATGGTCCG
CTGGAAGGCGGGCCTCGCCGAAGCCACGGCGAAGTGGCTGCTCACTCAGTTGGAAGACCGATGA

Protein sequence: (SEQ ID NO: 191)

VAPLDNAPPSGPLQDPSLARYSERQLAVANTWATHFSLAGTARTKFI RHYLRSTSTTRCWCITVAADNGVRYTIMRAGPLLQV
FDGQLIGAWECKPAHRI PASTPSRAGALKLLQRLQKFDDAVAVLSSYTKRAHDLATQMARDLGLQHRLLVYP SHSNKRY YAPR
HQFYLKQIGAVLRTFRQVLDQDLLFAIRS VRCLSPQLYNWLAQGDQVRRLLQMLKAQPVLTPLLVDC EGVWPHTTTNDNGESI
RHYLP CFPFQLD SERPQAAAMP CDLYLDMGRILGQVADEGISVINFFAWLFQAPRASIRFLSHVSPGRAGGALFHRKREGRHS
GWHALLLAASLGNRRPITRAQWTA FYAAYNAIPWQVHNAPDYNRLFNCGPSDWQDPAWLAITARLRDIKEFYTALDQNSQV
VRQARSALKAYLGHCTYRQAGNLVDDYHQVQRELRAAVQSSLPDLVDTD EYTTWEGMLSVGLIDCPNGLQIVELRCPADLYAE
HIALAH CIDSYDQAA YRGDCRLLSVREAGRPLASAELELRREHGEP IGRPWSPKHLSTVQLREFDNAPVPTDSPAGQAYRWF
ERIRSGAIATNLNWPDMTVHMTFRFANGRWKAGLAEATAKWL LTQLEDR.

RL091

DNA sequence: (SEQ ID NO: 70)

ATGCGAAAAGAGAATATATCTGCCGAAATCACAGAGCGAGCTTTTGATTTTTTCTATTGGTTCTCGCGATTGAGTTCAGCCT
CAAAGAGAATGGCTACTTAAAAAATTACAAACCTGGAGCTAGGGCAGAGCCGGGATGGGAAAATTTGTACAAAACCATTTCTG
ACAAATACTCTCTTTCCCAATCAGCCACAGCACTAATCGAGCAGAGTCCAGAGCAACAAATAGTCTGCCCGGTAGAGAGCTG
GGTTGGCGTCCGTTAAATTAGATGAGGACAAAAGCGACTTAGCTAGAGTCGCTCGCTTACTTAAGACCGTGCGAAACAATCT
ATTTACGGAGGCAAGCATGGTGGTGCCAACTGGGACAACCCAGCGAGGACAATACATCTTATCTTTTAAGTAAAGCTATCC
TTGACGAGTTTGTCTGCACTAGGAGACTTTGAGGCTGACTACAAGAGAATTTACTGA

Protein sequence: (SEQ ID NO: 192)

MRKENISAEITERAFDFYWF SRFEFS LKENG YLKNYKPGARAEPGWENFVQNHSDKYSLSQSATALIEQSPEQQIVLP GREL
GWRPVK LDEDKS DLARVARLLKTVRN NLFHGGK HGGANWDN PARTIHLILLSKAILDEFAALGDFEADYKRIY.

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND
PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September 12, 2003 Serial No.: Not Yet Assigned

Page 59 of 118

Customer No.: 21559

60/118
FIG. 300

RL092

DNA sequence: (SEQ ID NO: 71)

ATGCACATCGTAATCATTGAAGCCCCGGGCAAGCTGAAAAAGCTGAGGTCCCTTCTCCCTCGATTCTGCCGACGTGACCTG
 GCAGGTGAGGCGACAGCCGGCCACATCAGAGACCTACCCGTTACGGGCGAGGATCCGAGATGCTACCGTGGCGTGGGCC
 AGGATTTCAAACCGCACTACCAGATCCTCTCGGGCAAGGAAAAACCGTTCGACGCTGAAGGAGCTGCGGCAGAAAGCCGTG
 GAAATCTACGTGCGATCGGACCCGGATCGCGAAGGCGAAAGCATTGGCTGGCACATCCTCCAAGCTGCCGGGATCAAGAACTA
 CAAGCGCGTTGCCTTCAAAGAAATCACAAAGTCATGCATCACCGCGAACTCAGCTCGCCGCGTCCCTGGACCTCCCGAAGG
 TCGCCTCGCAGGAATGCCGTGCGGTATCGATCGCCTGGTGGGTATCTGGTCACGCCAGAGTTGCGGCGCGTGATGGGTAGG
 CCGACCACCGCCGGGCGGTGAGTCCGTGCGGTGTACCTGGTGGTCTGCGAGAGCGGGAGATCCGCGCCTTCACAGCAAT
 CAAGCACTTCGGGGTGAAGTACCTTCGTTTCGCCCAGCGACGGCCGTACCTGGACGGCGGAATGGGATCCAGTGCCCGTGT
 TTGCCAGCGAGGAGTTCCTCGTATGTCCAGGATCGTCAACTCGCAGAACTGGTGGGGCTATACGTAATGTCATCGTCGAGACC
 TGCATTGATAGCGAAGAAACCGATGCGCCTCCGGCACGTTTCATCTCCTCGCTCCAGATGGCCGCGGGAATCGCTGAA
 GTGGTCAACCCGACAAGACGATGAAGGTGCGCCAGCGGCTGTATGAACAGGGGCTCATCACTACCACCGGACGACAACCCCA
 ATATCTCGAAGGACTCGATGCCGGATATCCGTGCTGTGCGCAAAGCCTTGGGGCTGAAGTGTGTTGAGCAACAGCGGATGTTT
 AAAGCGGACCAAGACGCCAGGAAGGCCACCCGCCATCACCCCTACCGACTGGATGGCCGCTGCCGCGGTGAAACTGCTGA
 TGAGCAGGCGCTGTACAGCTCATTCGAGTCCGCGCGCTTGCCAGTCAAGCTGCGGTGTACGAGTGAGAACCATCA
 CCCTCCTGGGCGTCGGCCCCGACAAAAGCCGCTGCGCTTCGGCGCCAAAGGGAAGCTGTTGAACGTGCTGGCTGGAGAAAA
 CTGCTGCAGGGTGTATGACGCCGAGGAGCAGAAGAACGAAACGCTTCAAACCCCATCCCGATCCCGGCGCTGGAGCCACGCCA
 GATACTCAAGGTCTACAGCGGCGAGGTCTTGAGAAGAAACACCCCTCCCAAGCGATTACCGACGCCAGCCTGGTGGGCG
 AGATGAAGCGCCGCGGATTGGTGGCCATCCTCCTACGCTCGATCGTGAAGAACATCATCGACAAGGGCCAGGTGCAGATG
 AAGGGGCGAAGCCTGATCCCGGCGAGCTGGGAGAGGCCACCATCGCGCTCCTGGAGCACAACTTCAGTTCTCAGCCTCGA
 CTTACCCGCAACCTCGAGGTGCGCTTGGACCGGATCGCCAAACAGCGAGGACACCTACATGAACGTGGTCCAGCAGTTCTACC
 AGCTACTACAGACAGAGCTGCAGACACTCCGCGCGCTCCCGAGCGACAGGACGAACCACGCGCAAGCTCCACCGCCAGTATC
 TCCTCGGCGCCGACCAGCACTTCCTTTGCGGCAAGTGCAGTCTGCCCCTGGTTACCGCAAGAAAGCCGCAAGGCGGCTT
 CGACTTCTGGGGTTGACGCGGCTATCGAAACAGGGTGCAAGGTTAGTACCCACCAAGAGCGGCCGCGCTGACTTCGACA
 ACCCGCGCGGGCTATAG

Protein sequence: (SEQ ID NO: 193)

MHIVIEAPGKLKLRSLPSIRPDVTWQVEATAGHIRDLPHVHGQDPQMLTVGVGQDFKPHYQILSGKEKTVARLKELRQKAV
 EIVYASDPDREGESIGWHILQAAGIKNYKRVAFKEITKSCITAEISSPRRLDLPKVASQECRRVIDRLVGYLVTPELRRVMGR
 PTTAGRVQSVAVYLVLREREIRAFATAIKHFGVELTFVSPSDGRTWTAEWDPVPVFASEEFPYVQDRQLAELVGAIKRVIVET
 CIDSEETDAPPAPFISSSLQMAAGNALKWSFDKTMKVAQRLYEQGLITYHRTDNPINISKDSMPDIRAVAKALGLKCVEQQRMF
 KADQDAQEGHPAITPTDWMAAAAGETADEQALYQLIRVRALASQIEAAVYAVRTITLLGVGPDKKPLRFAGAKGLLNVPGRWK
 LLQGDAAEEQKNETPSNPIPIPALEPRQILKVYSGEVLEKKTTPPKRFTDASLVGEMKRRGIGRPSSYASIVKNIIDKGQVQM
 KGRSLIPGELGEATIALLEHNFSFLSLDFTRNLEVALDRIANSEDTYMNVVQQFYQLLQTELQTLRALPSAQDEPRASSTASI
 SSAPTSDFLCGKGLPLVHRKKAGKGGFDWFGCSGYRTTGCKVSYPTKSGRPDFDNPRGL.

RL093

DNA sequence: (SEQ ID NO: 72)

ATGGATCAAAGCCTTTGCACATGCATGCCAACGCCAATCGTCAACCCCAAGGAGCTGCGACTGTGCCACATGTTAGTCGGTAG
 AACTTTCCCGATAACATTGATCGCAGGCGACCATTGGTTGAGCTATGACGGCAGCGCTGGTGGGTCCATGCGGATGAGCCCC
 CGACGGAGGACGAGGTGGCGGCTCTGTTGGTCAAGGCTGGTGGTGTCACTACGTGCTGGTGCGGATAG

Protein sequence: (SEQ ID NO: 194)

MDQSLCTCMPTPIVNPKELRRLCHMLVGRTFPITLIAGDHWLSYDGSAAWVDADEPATEDEVAALLVKAGGVTTWCWG.

RL094

DNA sequence: (SEQ ID NO: 73)

GTGGCAAGGGCTTCCGAATCGGAAATCTCGACCAGTACGAGGTGAGTGTGTCAAAGAGAGCGACCGATACCGACAAGCTGGA
 CAGACGACACTTCAACGATCCCCACCGGACTGTACGGGCTATTGGTGTGAGGCGCGCGGAAAGGGCTACGGGTGTTTCGACT
 GCCCTACAGTCATCCTGCGATGCGGGCTCCTGGTTGAAAGGTTTGGCCAGGAGCAGCAACAGCAGCTCGACTTCTGA

Protein sequence: (SEQ ID NO: 195)

VARASESEISTSTRCSVSKRATDTDKLDLRHFNDPHRTVRAIGAEAAKGLRVFDCPYSHPAMRASWLKGFQAEQQQLDF

61/118
FIG. 30P

RL095

DNA sequence: (SEQ ID NO: 74)

ATGGCTACCCCGCTCTTCTGGGAAGCCAACATTGGCTCGGCGCCGGAGCACCGCAGCTTCCCCAACGGCAACAATCCCCCGCG
GCAGTTGCTGCGACTGAACGTGATGTTTCGACAACCTCGATTCCCGATGGCCAAGGTGGCTACAAGGATCGCGCGGCTTCTGGT
GCAGCGTCGAATGGTGGCATCAGGATGCCAGCGCTTCGCCGAAGTGTTCACGAAAGGTATGCGCGTCAAGGTGGAAGGCAGG
GCCATTATGGACCGCTGGCCGGACAAAGAGTCAGGCGAAGAAGTCCAGGCGCTGAAGGTGGAAGCCTCGCGCATTTCCATCCT
TCCGCATCGCCTGGCCGAGGTCAACCTGTTGCCAAGCCAGCATCAACAGTCTCGGAACGTCCCGCAGCAACCTGCTCAGCAAG
ATGCGCAATCGCAGCAGGACTACGACAGCGCTTCGACGACGACATCCCCATGTA

Protein sequence: (SEQ ID NO: 196)

MATPVFWEANIGSAPEHRSPNGNNPPRQLRLNVMFDNSIPDGQGGYKDRGGFWCSVEWWHQDAQRFAELFTKGMRVKVEGR
AIMDRWPKDESGEVQALKVEASRISILPHRLAEVTLTPQHQQSRNVPPQPAQQDAQSQDYDSAFDDDDIPM

RL096

DNA sequence (SEQ ID NO: 75)

ATGCGCGAGCTCGATAAGGACCAGCAAGGCGCTCTGGAACAAAGTGCCTTCCGCCACTGCAACAACTGCCTTCCAGGCGCT
GCAACACAGTGCCTCACTAAAAGGCCTTTTAAAGCCTTTTAAAGGTAATAGGGAGCTGGCCAGTTGGCGGAACAGTGCGAAG
CCATGGAGCAGGGATTGCTTGAACCTGCCCAGGACTGCTGGCCAGGTTTCGTCGCCACCCTTCACTCTACTGCCACCCGA
CTCATCGAGCAGCGCACATCCGCCCGCACACCTTTCTCCGCTGGCAGCACATTGCATCCCGGCGGATGGGCGTCGGGGTGTG
GACGGAAATGCTGCGCCAGGACAAGACCCCGGAATACCTGCTGCAAGACCTCTACGAGATGGAGCTGCAGCGCATCACCTCA
ACATGCAGATCAGCCTGATCCACTCCATCGGCAAGCAGGCCGCCGAGTGC CGCGGAAAAGATGGGCCAGGCCGAGGCCGAGTTC
ATGGGCCGACTGCAGCAGAGCACCAACCACCACTGA

Protein sequence: (SEQ ID NO: 197)

MRQLDKDQGALEQSAFRPLQQTAFQALQHSASLKGLLKPFKGNRELAQLAEQCEAMEQGLLELAQGLLAQVRRPPTLLPTR
LIEQRTSARTTFLRWQHIASRRMGVGVWTEMLRQDKTPEYLLQDLYEMELQRITLNMQISLIHSIGKQAAECAEKMGAEEF
MGRLLQSTNHH

RL097

DNA sequence: (SEQ ID NO: 76)

ATGGCTGAAACCCATCGGCTTCAGATCGGCTCTCTCCGCAGCGATGTGCTCTGACGCTTCACACCTATCACGCCGCCCGCAT
CTGGACCGGCCGAGAACTCGGATGCCAAGCACAGCATCCTCGGCCTCTCCGGCTTCTGCGCATACGTGAATCGCATGCACC
GCGGGGCGAGCACAGGACGATCCGTACTCCGACTGGTGGCTGGTTCAGATCGAAGAGAAAGTCGAGAGCTGCCAAGCCGCACTC
GAGGCCATCGACCAGCGACTGGATGACGTGATGCCAAGCTGCCCGGACCCTCGATATCTCCGAGAACCTGTCCGTTACACC
GGTCAAGGTCCCGTTGTTTCATCTCCAACCTCTCGGCTTCAAGGCGAGTCTATCTCCTGACCAACTACGACGAACTCGCCCGTC
GAATCTGCTGGCCAGCACGTGCGGCTGGTGGTGGCGGACATGGAGGTCTGGCTCGACGAAGGTGCGTGGTGTGCTGCGA
AGCCTGTTTGGTCTGGCCAGAGCTACCACTTCTCGGGCGCCACTCGCGACGACTTCGCCGCAACAATGCTCGCGCCGAAGC
CGCGCGGAAGATGTACGAGAAGTTCGGCGAGATCCCGCAGGACATCCTGGAGGGCACTCGACGCTCGAACTTCGCTCCGCCGA
TCACCCGGGGCCGCTCTGACGGTGATGCCGATGATGACGCTGACCGTGTGCAACTCGAGGACTGA

Protein sequence: (SEQ ID NO: 198)

MAETHRLQIGSLRSDVALTLHTYHAARIWTGRQKSDAKHSILGLSGFCAYVNMHRGAAQDDPYSDWWLVQIEEKVESCQAAL
EAIDQRLDDVMAKLPAITLDISENLSVTPVKVPLFISNPLGFKAVYLLTNYDELARRILLAQHVGLVGRRDMEVWLDGASVLR
SLFGLAQSYQFSGATRDDFAANNARAEEAARKMYEKFGEIPQDILEGTRRSNFAPPITRGRSDGDADDDADRVELED

RL098

DNA sequence: (SEQ ID NO: 77)

ATGTTCTTGAGCATGGCCCTTTCTTTTGGTGGTCTCGTTCGTTCTTCTGCACTTTTACAGATGCGTGGAACGACCGAGAACT
CAGGCTGTTGTTAATGCTGATCGTGTTCGGGTATTAGTAACCGTGTGACCATACGGTTGAGATGTATCGCTTTGAAATGG
CGGAAAAGCGATGTGGGAGCTTTATGCAACAAAGCCAACACTACATGAACTGCCAACCAGATTACCAACGGTAG

Protein sequence: (SEQ ID NO: 199)

MFLSMAPFFLVVLVLSALFTDAWNRDLRLLLMLIVFGYSVTVLTITVEMYRFEMAEEKAMWGALCNKANYMNCQPDYQR.

62/118
FIG. 30Q

RL099

DNA sequence: (SEQ ID NO: 78)

ATGAGAAAGTCTCGGTCGGGCGTCGTGTTTTTTGGTGATGCGGCCCGCATCACTCTCCAGGTCCTGACCTCCGCGCCGCGG
CGAGCTGGGTGATTCCACTGGAATCACTCCACCAGGAGCCGACCTCCGCGCCGCGGCGAGCTGGGTGATTCCACTGGAATCA
CTCTGCCAGGGATCCACTTCGGTATCGGCGGCAAGATGGGTGTTTCGGGCCGAAACACTTCGCCAAAGCGAGGCATCACCCT
CACGAGGAACCAACAATGTTCTTGA

Protein sequence: (SEQ ID NO: 200)

MRKRSRGVVFFGDAARITLPGPDLRAAGELGDSTGITPPGADLRAAGELGDSTGITLPGIHFGIGGKMGVSGRNTSPKRGIT
HEELKQCS.

RL100

DNA sequence (SEQ ID NO: 79)

ATGAGGCTGTGCGGCTTCCCATTTCGACACTGCTGGACTCGGCCTCGGGGCATCTCGAGGCCCATTTGTATAAGAAGCGGCT
TGCTGCCGAAAGCGCGGAACCGCTGGCTCAACAATATCCGGCATCATTTTCAGCGGAATCCTCATGAAACCGTTCCACGGC
GCCTCCTCCTGGATAAGCGTCTTACTCCGCTGGAGCGGAACCTGCTGGCAAGTTTCCGCTTGCTCATCAACGACGACGGACTG
ACCGCGTTCCCGACATACGAGCAACTGCGCCCTATCTCGGTATGCAGCCGGGCAAGATCGCCTCGCGCGAAACCATCGCCAA
GGCACTCACGGTCCTTCGTCTGACCCGCTGGCTCAGCCTCGGCCGACGCCTGCGCAACGACCTCAACGGACAGGTCCAGGGCA
ACGTTTACATCCTTCACGACGAGCCTGTCTCTCCAGCCGAAGCCTTGGAGCTGGACACCGACTACATGCAGTTGCTGAGCCAA
TCCACCGGTACGCGCAACCGAGCCATACGCGAAATCGGGCAGATCATCTGGCGGGAGTTTCAGGGATGATCCGGACGTGGTTCG
CCGCTCCCTACCCATCTGGAGAAGCTCGAGGGACGCTTGAACCAACAGCAATGGGCTATCGATAGTCAGCTCGAAGCGGATC
CAGCGGCAGAGTTCGGCATCCGAACCTCTGTCGGATTTACCTCATTCACCCCGAGTTTCGGATGCCGAACCTAGTGAAATCAGC
GGCAAGCAATGCGCTCTACCGCTGAGTTCGGATACCGAACCCGACAGAATCCGCCGAGTACGCCCTTGGTTTCGGATGCCGAA
CTCATATAGTACGTATACATACAAACAAGATTCTGTATGTA AAAAGCCAGTACAACCGCGAGCACGCCGAGGAAGCCCATCCGA
ACTGGCAGGATCTCCTGCACGCACTGGAGGCCGAGCAACGGATCCAGGCAGTAAGCGCGCTCAGACGGGTGTCCGAGGATCTT
CGGCTACCCATCATCGAGCAGTGGCAGCACCGTTGTGCGCGCGGAACAGTCAGCAATCCGTTCCGGCTACCTCATGACGCTCAT
CCAGCGTGCAGTCCAGGGCAAGTTCAACGCTTCTTGGGCTCCGGAAGAACCGGCTGAGCGAACCATCCCGCAACGGAACGCC
CCATTCGTGCTCCGGCACCATCAAGCCCCATAGCGCCTACACAGCCTCAGGTCCAGCCCCGGGGGATACCCGGACAGGGAGC
GAGGTCTCAGCCGGCTCAAGGACCTCATTCGGCCCCAGGCACGGATCGAGCGTGCCATCCGAGCGGGGTGATGATTTCATGA

Protein sequence: (SEQ ID NO: 201)

MRLSRFPFISTLLDSASGHLEAHLYKKRLAESGEPLAQYSGIIFSGNPHEVTPRRLLLDKRLTPLERNCWQVFRLLINDDGL
TAFPTYEQLRYPYLGMPGKIASRETIKALTVLRLRWLSLGRRLRNDLNGQVQGVNYILHDEPVS PAEAELELDTDYMQLLSQ
STGHGNRAIREIGQIIWREFRDPDPVGRRLPHTLEKLEGRNLHQQWAIDSQLEADPAAEFGIRTLSDLPSTPSSDAELSEIS
GKQCALPLSSDTEPRQNPSTPLVRMPNSYSTYTYKQDSVCKKPVQPRAREEAHPNWQDLLHALEAEQRIQAVSALRRVSEDL
RLPIIEQWQHRCAGGTVSNPFGYLMTLIQRVQKGKFNASWAPEEPAERTIPATERPIRAPAPSSPIAPTQPVQPRGDTRTGS
EVLRLKDLIRPRHGSSVPSERGDS.

RL101

DNA sequence: (SEQ ID NO: 80)

ATGTCGAAGTCGACGATCAATGAAGCGGTCTGACGAGGTGCTCAACCACCTGCGCAACGGCCAGCTCAGGCGTTGTGCCGA
GATGGGGCTGCGGCCGAGATTCTGGCTCAGCTCCAACAGCCTGCCGTATGAGCATCCTGACCAATACCCCGGTTTCTGGG
TAGATGTCAGAGTGAACATCGACGTATGGAGAAAATCCTGGCCACAGCCGAGCGCAGCGCGAGGAAGACCTGCAGATCGAA
CGCGCACTGAAGCTGGGAGCCACCACAACGATGATCCAGAGCTTTTTCGGTCTGTGCGCGGAGGACACCGCCACCAAGCGCTT
GATGCTGGAGATCCACCCGCGCCGCGTCTGGCGGCAGCTCGATGAACAGATCGAGCGCCAGATATGGTTCCGCTGGGAGC
ACCTGATGCAGGAAAATCAGGTCCGCCTTGAAGACAGCATGGAGTTGCTGGACATCGCGATGATCCTCACAGAGGAAATCAAC
GCCGAATCGAACAAGACAGTCCAGAATTCATCAGCCTCGCCATTGTTTGGTCTCTCATCCAGAGCTGGTTGAAAGACGGGCT
CTATCCGTCTGGCAAATCGAGCCAGAGCCAGCGGGCCTGCAAAAGTCCCAATCCACTCTTTACCTCGCTAGCGTCAGCTCAC
ACCTGCCCCACTCTGCCCATCCGCAACAACGCAGGTGAACGCTGAGACAGAACGTCAACAACACTGAACTGGTTCACTCG
GAAGCGGACACAGCACCATGA

Protein sequence: (SEQ ID NO: 202)

MSKSTINEAVLTQVLNHLRNGQLRRC AEMGLRPEILAQLOQPAVMSILTNPVSWVDVRVNIDVMEKILATAERSAQEDLQIE
RALKLGATTTMIQSFFLSPEDTATKRLMLEIHPRRGRWRQLDEQIERQIWFREHLMQENQVRLED SMELLDIAMILTEBIN
AGIEQDSPEFISLAIWLSLIQSWLKDGLYPSGKSSQSAGLQKSQSTLYLASVSSHLP HSAPSATTQVNAETERQQLNLVQS
EGDTAP.

63/118
FIG. 30R

RL102

DNA sequence: (SEQ ID NO: 81)

ATGAGTATGGCCAAGATCAACCCGCAAGATCTGAAAGACCGGCTACTTGCCCTGGTTTTACCGCACCGCCCAAGGTTCTGGA
GCAGTTGTCCGACCCGATCAGCGACACGCCATGAGGCTGACACTCCACGACGCTCTCCCTGGCAGCATAACCCCGGACCA
CTCGAAACCCGAAATACGATGAGCTGAAAGAATCGATCCGACATCGAGGCCTCGATACGCCGCCACCGAGTGAAGTCCGCT
GGAGAGGACAAATACCGCATTGCAACGGCGGAAACACTCGCTGGAATTTCTAACGACCTCTACAAAGAGACCGGAGACGA
GCGCTATTTTACGCTTCGACTGCCTGTTCAAGCCCTGGGACAAGCAGCGCGCGAAATCATCGCGTGAACCGGTCAATTTGGCCG
AGAACGATCTGAAGGGCGACCTCAAGTTCATCGAGCGCGCGTTGGGGTGCAGAAGGCGAAATTTCTTTACGAACAAGAGAAC
GGCGGTGAAAGCATTTCCAGCGGAGTTGGCAGCTCGGCTAAAAGCGGACGGCTACCTGTATCTCAATCCCATATCAGTAA
GATGTTAGACACTATTGAGGTATTGGCGCGCGGCGATTCTGTGATGCTGTATTAGGGCTCGGTAAACCGCAAATCGAGAAAC
TCCTGTCACTCAGAAAGTGGCATCTCTGCTGGGCACGCTATACGCTGGTGAAGGGGTTGACTTCGAAATGCTGTTCCAG
GACACCTGGCAATCTTCGACAGTAGCCCTGACGAATTCATTTTCGAGCGTTTCCAGGACGAACCTCATCGACCAAATGAAGCG
CCCCCTGGGCCTGCGTTATGACCAAATCCTGCTCGAGATTACCAACGGGCAGCAGGAGCAACGCCCGCGCACTCTGGTCGACC
TGCCACACCTGCCGCACCACTCACTCCACCAATTGGGAGGAAAACCTGCTGCGTCTGCTACTGGACAAGCACAACAA
CAGAGCCCCCGCGCAGATCCCCAAACGTCAGGACAAGGAGCAACCCGGTAAATCCCCCTCCCCCGCGCGCTCCGCCACCA
TGTCACAAAAGCAATTGCCCCGATGAGGAGCGTGGCGCGTCTTGGCAGGCCATATCGTGAGCCCGGTATCGACTAAGATCC
AGCAGACTCGCCAACCGCTGGCCGGCCTCGAGGGGGAACATCTACCTGTCTTCGATGAAAACAGCTTCGAGGCAATCCAGTG
CAAGTCGGTGGCCTGCACCCGATCACCGATCTCTGGTACATCGAGCGGTTCGATCGATACCCCCGAGATCTGCGACAGCACAT
CGCTGATCTGGCTGAAGAGATCGCTCTGCATGTGGCGCCCCAGGCGAGATCGTCAGGATTAGGGCGGTGTGGGTTACAGT
ATCGCGAGCCCAATGAAGACCATGAGATTACTGATTACGCGCTGCACCTCATGACGCTGCTTCAAGCGGTGAGCGGGCAGGT
CAAGTCGTTCTGAACACTCAGCATCAACAGACCTGCCGCGATGCACTGGGTGAATTCCAGTTCCTAGCTGGCCTCGCTCAGTT
GCTGCTGGGCCAACCCACCACAAGTGACAAGCCATCTGCGCAGGAGCCGCTCAATGACGAAGCCCTGGTGAAACTGTTC
GGATCATTCGTCTTGCCCGACGCTGGTTGACCTTGAGCTGCCCGCGCGCTCCGAGCAAGCAGCTACTGACCAGTGA

Protein sequence: (SEQ ID NO: 203)

MSMAKINPQDLKDRLLAPGFTAPPKVLEQLSDPISDTPMRLTLHDVLPWHDNPRTRNPKYDELKESIRHRGLDTPPPVTRRP
GEDKYRIRNGNTRLEILNDLYKETGDERYFSFDCLFKPWDKQRGEIIALTGHLENDLKGDLKFIERAVGVQKAKFLYEQEN
GGESISQRELARRLKADGYPVSQSHISKMLDTIEVLAPAIPVMLYSGLGKPKQIEKLLSLRKSASSCWARLYAGEGVDFEMLFQ
DTLAI FDS SPDEFI FEFQDELIDQMKRPLGLRYDQILLETNGQQEQRRGTLVDLPTPAAPPQLPPIGOENPAASSTGQAQT
QSPAADPQTSRTRSNPNPLPPPAPPPVQKQLPDEERAVALGHIVSPVSTKIQQTRQRLAGLEGEHLVPVDETALQAI PV
QVGGLHPITDLWYIERSIDTPEILRQHIADLAEBEIALHVGAPEIVRIQGGVGYTYREPNEDEHITDSALHMLTLLQAVSGQV
QVVLNTHDQQTCDALGEFQFSAGLAQLLLQPTTSDKPSQAGRLNDEALVKLFRIIRLARRLVDELPPAASEQAATDQ.

RL103

DNA sequence: (SEQ ID NO: 82)

ATGACCATGGCCCGAGAAACCGAAGATAAGTTGCTTGCTCCGTATGCCCTTGGGCTTGCGCGATCAGCTAAAGCAAAAAGCCGC
GGATAACCACCGTTCCGCCAACAGCGAGATCGTCTACCGACTGGAGCGCAGCAACGCGCTCGAAGAAGAACTCGCGCGAGCAA
ACCGAATGGTTCGACGAACCTCTTCGCCAAGAACCAGCGCCTGCAGGCTGAGCTGGCGGCGGCGAACACGCTCAGGTGGCGGAG
GCATGA

Protein sequence: (SEQ ID NO: 204)

MTMARETEDKFVVRMPLGLRDQLKQKADNHRSEANSEIVYRLERSNALEEELARANRMVDELFAKNQRLQAEALAAANTPQVAE
A

RL104

DNA sequence: (SEQ ID NO: 83)

ATGCCTATCAAACACGCCATCGTCCACCTGATCGAGAAGAAGCCCGACCGCACCCCGCGCTGCTCCACGCGCGCGACGCCGA
GCTGGGCGACTCCCAGGCCATCGAGAACCTGCTGGCCGACCTCAACGAAAGCTACAACGCCAAGAACAAGGCCTGGGGCTTCT
TCCAGGGCGAGTCCGGGGCCTACCCGTTACGCGGCTGGCTCGCGGAGTACCTGGAGGGCGACCCGACTTCGTGCGCTTCAGC
CGCGAAGCGGTGAGCACCTGCAAAAGCTGATGGAGGAGTCCAATCTCTTACCAGCGCGCCACGTCTGTTCGCCCCTACCA
GCAAGGCATGACCGACTACCTGGCGATCGCCCTGCTGCACCACAGCGAAGGCGTGGCGGTGAACGAGTCTGCTGGAGGTACCCC
CGTCGCGCCACCTGGACCTCGGCCAGTTGCACCTGGCCGCGCGGATCAACATTTCCGAATGGCGCAACAACAAGCAGTCAAG
CAGTACATCTCGTTTCATCAAGGGCAAGGGCGGAGGAGTCTCCGACTATTTCCGCGACTTCATCGGTGCGAGGAAGGGGT
GGATTTCGCGGAGCGAGACGCGCACCTGCTGAAAGCCTTCAGCGATTTCTGGAAGAGCGAGGACATGGCCGAGGAACAGGCC
GCGAGAAGACCGAGACGCTGGTCGACTACGCCACCTCGCAGGCGCGCATCGGCGAGCCGATGACCTCGACGCGCTTTTCGGA
CTGATGGACGACAGCAACCGCGGGCGTTCTACGACTACATCCGTAACAAGGACTACGGCCTGTGCGCGGAAATCCCGCGCGA
CAAGCGCACCTCAACCAGTTCGCGCGCTTACCAGCGCGCGCGAAGGCCTGTGATCAGCTTCGAGGCGCACCTGCTGGGCT
CCAGGATCGAGTACGACGAGGAGCGACACGCTGCAGATCAGCAGCCTCCCACTCAACTCCGCGACAGCTCAAGCGCGCG
AAGGCCCAAATTGGAGAATGA

64/118
FIG. 30S

Protein sequence: (SEQ ID NO: 205)

MPIKHAIVHLIEKKPDGTPAVLHARDAELGDSQAIENLLADLNESYNAKNKAWGFFQGESGAYPFSGWLGEYLEGDRDFVGF
REAVEHLQKLMEESNLFTGGHVLFAHYQQGMTDYLAIALHHSSEGVAVNESLEVTPSRHLDLQQLHLAARINISEWRNNKQSK
QYISFIKGGGRKVSDFRDFIGCQEGVDSPPSETRTLKAFSDFVESEDMAEEQAREKTETLVDYATSQARIGEPMTLDALSE
LMDDQPPRAFYDIIRNKDYGLSPEIPADKRTLNQFRRTGRAEGLSISFEAHLGSRIEYDEERDTLQISSLPTQLRDQLKRR
KAQIGE.

RL105

DNA sequence: (SEQ ID NO: 84)

ATGCGTAGTTTCCTTCGCGGCGCCCGGAAAGCGTTCGCCGGCTGGTGGCCTTCGCTCAAGCAGAAGGCTGGAGCGTCGACCG
CTCCGACGGCGGCCACTTGAAGCTCAGCAAGATCGGCTGCGCCTCGATCTTCATTTCTTCCACGCCAAGCGACGCACGCGGCG
AGCTCAATGCCGCGCCCTGCTCCGTCGAGCCGACAGGCAGCGTTCCTGAACCAGGAGTCTTTCTGA

Protein sequence: (SEQ ID NO: 206)

MRSFLRGARESVRRLVAFQAEGWSVDRSAGGHLKLSKIGCASIFISSTPSDARGELNARALLRRADRQRSLNQESF

RL106

DNA Sequence: (SEQ ID NO: 85)

ATGCCTGACGTCACAGCCTACCGGCCGCTCGAGCACTTCCAGAAAGTCGAGCTGATGCTTGAGCTCAAGTTGCGTGAAGGTCC
TTCGTGGATCTGTCTGAACTGCGGCTATCACCTGGATGGCAGCGGCGCACAGCCCTGCCCTGACTGCGGAAAGTCGCGCTACT
GGACACGCGGTTGGAGTGTAGGTCGTGGCCATCGCTTCTCGGCAGCAAGGGAAGAGTGGGAAAACCGCCTCAGGACACGGTCC
CGGTCACCTGTGCGCTCAACGGCACCACTGACAACTGACGAGCTATGCACTCAACTGCGCACAGAGGTCCGCGATGCTGCGTTC
CGCGCATGACGACCTGGCCTGCAGCCGGCAGAGCGATCGTCGAGCCTTCAGGCGCTGGTGAACCGTCTCCTGGATGCCGCCG
CCACCGATAGCCTTCCCGCTCCCTTGCAGAGATGGAGACCTGGCTGCAGCTCAACAGCGAGGAGACCACGAATGCGTAG

Protein sequence: (SEQ ID NO: 207)

MPDVTAYRPLEHFQKVELMLELKLREGPSWICLNCGYHLDGSGAQPCPDGKSRWYTSWGSVGRGHRFSAAREEWENRLRTRS
RSPVASTAPVATDDVCTQLRTEVRMLRSAHDDLACSRQSDRRSLQALVKRLDAAATDSLPRSLAEMETWLQLNSEETTNA.

RL107

DNA sequence: (SEQ ID NO: 86)

ATGAAGGCGTCCCAGACCTATCAGTGCATCGTCAAGTTCGATGGCGCCGGTTTCTGGACCAATACCATTCAGAAGCAGCGTGC
GACCTGCACCTGGAGCGACAAGGTGGCAGCCTCCCGCCTTGCCTGAACGACTGTTTGGCGAGGACACGCATACATCACCCTGA
TGCCCGGTACAGGCAGGCGACACGAAAGCGCATCGAGAGCCGCTGGGCGCTGTCTGTAGAAATCCAAGGAGGTAGCGCGC
GATGCCTGA

Protein sequence (SEQ ID NO: 208)

MKASQTYQCIVKFDGAGFWNTIQKQRATCTWSDKVAASRLAERLFGEDNAYITRMPVQAGDHEKRIESRWALSCRNPKEVAR
DA

RL108

DNA sequence: (SEQ ID NO: 87)

ATGAACACTGAAGCCCGCTTTCGAGTATCCAGCCTCGGCCGCGTTACCGACTCGGCAGTGGTTCATGCCAATCACGTGG
GGTCAACCCCATCGAGCTGGACGCCCTCAGCCAAGTGATCTCGCGCCTTTCGCGGGACGAGAGCACGGTCGCACCCAGTTCGA
TGGAGCGAGAGCTTCGTGAGCTGGAGGAAGTGGGGTACATCGAAATCTCGACCACCCAGGCCGGGACTCTGGTGGTCACTACG
CGCGCTCCGGGGCAATGTCTTTCGGCTTACTTCTGGTGGTATGGATCCCGCGACACCTGTTGAGCTGCTCGCTGAAAGTGAG
CCTGGTGCCGCACCTCTGCTGCGGCACTCAGGACTCCCAGCACCTCACCGCCGTGTTCCGCATTGCAGGCAGCAAGGACGCCG
CGCGGAGTTCTGTCATCAGTTGGCCAACAATATCCCGGGCATGAGCCGGAGTTGCCCGAAGTGGTGGCCGTTTCAAGTCCGT
GATGCACTCAGCAAGGAGGCCGAGTCATGA

Protein sequence: (SEQ ID NO: 209)

MNTEARFPSIHASAAFTDSAVVHANHVGVPNIELDALSQVISRLSRDESTVAPSSMERELRELEELGYIEISTTQAGTLVVTT
RAPGQLLSAYFWSVWI PRHLFSCSLKVS LVPHLCCGTQDSQHLTAVFRIAGSKDAAREFLHQLANNYPGHEPELPELVAVQVG
DALSKAEAS.

65/118
FIG. 30T

RL109

DNA sequence: (SEQ ID NO: 88)

GTTCCTGCAGATCAGCCGGGAGGACTCGAAGTGAGTGTCTTGGAACTTACGCCGCCGCACTCCGTCGAGGCGGAGCAAGGGGT
ACTCGGCGGCCTGATGCTGGACAACGCGGCATGGGACATTGTTCGGCGATCAGTTGCAGAAGGAGGACTTCTTCGGCATGAGC
ATCGCTGATCTTACCGCCATCAGCGAGTTGGCCGCGAAGGATGCTCCGTTTGATGTCGTGACTGTGTCGGAAGCGATCGAA
GACCTTCCAGAAGCTGGCGGGCTGGCCTACCTCGGCCAGCTCGCCGACAACACGCCCTCCGTGGCCAATATCGAGGCTTACGC
GCAGATCGTTTCGGATCGGGCACACCTGCGGCAGCTGATGTCTTCGGGCACCACTGCACCAGGACCGCCTCGAACCACCAGG
CAAATCCCTCTGAGGTTTCAGGAGGAGATTGAGCAGAAGCTGTTTCGGCCCTTGGCCAGGACCACCACAACGCCGATTTCTGTCGA
TATCAACAAGAGTCTCAGGAAGATCGTCGACACCATCGATTACCGCTTCAACAACAACGTGACGGTAACGGGGGTCCCGACTG
GCCTGAAGGATCTCGACGCACTCACCGGCGGACTACAGAAGTCGGATCTCATCATCGTCGGTGCCCGCCCGCGATGGGCAAA
ACGTCGTTTGGCCTCAACCTGGTCGACACCGCGCTCCAGAGCGACCAACAGAAGTCTGTTTCAGGTGTACAGCATGGAGATGCC
GGCAGAGCAGTTGCTGTTTCAGGCTTGGCGCCCTGTTTCGGCCACCTGGACCTGGGCAAGCTGATGAAGGGCCAACGCAAGAAG
AGGATTGGCCCAAGACTGTCTGGCGCGATCCAGCGCATAAACGACTATGGCAGCCGGCTGGTCATCAACGATCAGGGCAACCTC
ACGCCGACAGAGCTGCGCGCCAAGGTTTCGGCGGGCGGCCAGGAAGTACGGACACCCCGCGCTGATATTGGTCGACTACCTGCA
ACTGATGAGTTGGCCAGGCTGGAGAATCGAGCCACCGAGATCTCGGAAATCTCCGCTCGCTGAAAGCGCTGGCCAAAGGAGA
TGGACTGTCCCGTCGTAGCTCTATCCAGCTAAATCGCGGCTAGAGAACCAGGACGAACAAGCGACCGAAGTGGCGGACCTA
CGAGAGCGGCGCAATCGAGCAGGACGCGGAGTATCATGTTTCGTGTACCGGACGAGGTCTACCAACCCCAACACCGAGGC
CAAGGGCATCGCCGAAATCATCATCGGCAAGTATCGCAACGGTCCGATCGGCACCGTCCACACCGCCTTCATCGCCAACCAGA
CCCGCTTTGCGGACCTGGCGCGGGGACCTGGCAA

RL110

DNA sequence: (SEQ ID NO: 89)

ATGACTCGCTCTGCTCTCTCGACCATCGCCTACGAGGCCCTGGTGCGTGCCCGCGCAAATTCAGCAACCGAGAGGAGCGCTG
CATCCGCGAAACCTGGACCGCCGAACAGGAACTGGTGCTGTGCGCCTGTATCCGGATATGCCGAACGAGGTCTGGCAGCCA
GGTTGAACAAAACGCTCCAGCAGATCTGCTCCAGAGCGTATCGGCTCGGGCTGAAAAAAGCCCTGAGTTCTCCAAGAAGATC
CGGCAGGACTGGGGCAGCGCAACTCGGTTCAAGAAGGGAAACACCCCATGGAACCTGCGGCATGAAGGGGCTGCCCGCGGAGG
ACGCGCACCCAGAAACGAGTTCAAGAAGGGGCAAAAGCCCCACACATGGCTCCAGTCGGCAGCACGCGGGTCAGCGCTGATG
GCTACCTGCAACGAAAGATCTCGGATACCGGCTATCCCCCGGGGACTGGAAGGGCATCCACATCCTGCTCTGGGAAGAACAC
TTCGGCCCCATCCCAACCGGCCATTGCGTCTGCTTCAAGGACAACAACAGCAGAACGTGTCATCGACAACCTGGAGCTCAT
CACCCGGGCGCAACGCATGCGCCGCAACTCCATCCATCGCTATCCACCTGAGCTGAAGAGCGCAATCCGCGTTCATCAGCAAGC
TCAAACGCACCATTCAGGAGGTCGAGCATGAAGAACAAGATTGA

Protein sequence: (SEQ ID NO: 210)

MTRSALSTIAYEALVRARRKFSNREERCIRETWAEQELVLLRLYPDMPNEVLAARLNKTLQQICSRAYRLGLKKSPEFSKKI
RQDWGSATRFKKGNTPWNCGMKGLPARGRAPETQFKKGQKPHWTLPVGGSTRVSADGYLQRKISDTGYPPRDWKGIIHLLWEEH
FGPIPTGHCVCFKDNKNQNVVIDNLELITRAERMRRNSIHRYPPPELKSARVISKLRITIQEVEHEEQD

RL111

DNA sequence: (SEQ ID NO: 90)

ATGGACAAGCAAAAAGTCTCTCGCAAGGTGAGAAAGCTGATGGCCCTGGCGAATGCCAAGGGGGCCACGCCGAACGAGGCGGA
AACCGCATTGCGCCAGGCGCGATCTTGAAGCGGCAGTTTCGACCTCAGCGATGCGGAGATCTCGGCCACACGGTGGAAACCG
CGTGCGTTCCCACTCGAACCAGGCGCTCTCTGCCCCATGGCTGCATGAACCTGGCCGGGATCTGCGCCAGTTCTCTCGGCTGC
GACTACCTGGCGGCATACGCGATGCCAGCGGGCTGGACGTTCAAGTTTCATGGGCCGAGGGATCGGCCCTGAGCTGGCCGCTCA
CGCTACTCTACGCTCCACCACCAACTGGTGGCAGCGCTCGGCTCATGTGCGCCCAACAGAAGCGTGCAGCTGTGACCA
AGCGTCGTGCGAGCAAGCTCTTTCGTGCAAGGCTGGCTTCTCGCAGTGCCTTCGTGGTACGTGAATTGCTGGCAGGCGGGAC
GAGTCGACTCAAGCAGCCATCAAGGCCTACCTCGAACTACACCATCCGGCGTTGAAGTACCTGGAGCCGGCGGCGCTTACGAA
GGCCCTTGCTATGACCAGGCTCGCTGCAAGCAGGCTGGGAGCACGGCAAAACACTCGCCTGCACCGCGGTGTACGCCGCG
GAGTTCAGGGCGCGCTCGAGCAGGGAGGTTCCCAATGA

Protein sequence: (SEQ ID NO: 211)

MDKQKVLAKVEKLMALANAKGATPNEAETALRQAAILKRQFDLSDAEISAHTVETACVPTTRRSPAPWLHELAGICASSFGC
DYLAAYAMPAGWTFKFMGRGIGPELAHAHAYSTLHHQLVAARSAHVAQQKRKLSKRRRSKLFVEGWLLAVRSLVREFAGRPD
ESTQAAIKAYLELHHPALKYLEPAALTKALAYDQASLQAGWEHGKNTLRHGVSRVQGALEQGSQ

66/118
FIG. 30U

RL112

DNA sequence: (SEQ ID NO: 91)

ATGAGTGACCCCAAGCTCAAGCCCTGCCCGCTCTGCGGCAGCACGAACATTGCAATGCTGGAACCCGAGCTGCTCGACACC
GATGCCTGGAACCTGTGCCATTGAATGCCTGGACTGCCAGGTTACATCGGGCCGTCCTACTGCGAGCCAGACCCGGTAACA
GCGAGGTATTGAGCACAGATCGACTGGAATAGACGCCCAAGCGCAAAAAACCACGCGGACGAGCGTGAGCAGTTCTTGATG
GCCAACCTGCTCGCCGCCCTGGAGGTGCGACTGGGCGACGTAGCAGCCCTGGCTATTGTGATCGGGTAAGACAGGGCCACA
GACCGAATTTACCCAACTTCGAACCTCTCCCTGTTCCGCGAGGCTGGCTCGATGTACAGGCCGAGCGCCGGCGCCAGATC
ACCGTCGAAGGTTTCGATACCAGCAACGACGACGCTAGCGCTGGCCTGATCGCCCTGGCGGCCGGCTGCTACGCGCTCCAT
GCCGGCGGCATCGGCACCGACTGGCCGGGCGGCATTGCGAATGGCTCTGCACTGTTCTGGCCCTGGGACGAAGAGTGGTGG
AAGCCTAAGTCGGCGCGCGAGAACCTGGTACGCGCCGCGCCCTAGTGCTGGCCGAGATCGAGCGCTGGACCGCTCCGCC
ACCGAGCAGGGCTCAACCATCTGCAAGGGGGGCGCGTAA

Protein sequence: (SEQ ID NO: 212)

MSDPKLKPCPLCGSTNIRMLEPELLDTDWNCATIECLDCQVHIGPSYCEPDPVTARYSAQIDWNRRPSAKNHADEREQFLM
ANLLAALEVALGDVAALAIVDRVRQATDRIYPTSNLSPVPQAWLDVQAERRRQITVEGFDTSNDDASAGLIAAAGCYALH
AGGIGTDWPGGIRNGSALFWPWDEEWWKPKSARENLRVAGALVLAIEIRLDRSATEQGSTICKGGA

RL113

DNA sequence: (SEQ ID NO: 92)

ATGAACCTCCAGAACCGCAACAACCTCCTACTGAGCTTGATCGCCGAGACCCAGTTCGACGCCTACGTGCAAGGCTACATG
GCCAAAGCAGGCGCTGCCGCCGGTGCTTCCGAGAATCTGCAAATCGAGGCTGAAGGTGCTGCGATGTTGCAGGGCCTGGTC
GCTCCGGTTCGCGCTCAGCAGCGTGCCCTGTGGACAGTCCCTGCGAAGCGACTGCTCCAAATCGCCCACGACCTACTGTTG
CAGACGAAATCGCAACTGGCCATCGCGCCAATGCCAGTTGATCCAAAGTGATCCAGCGGGACATGAACAGGGCGATCTGG
AACATAGCTACTGCCATCGATCACCTGGCCGAGTTCGCCCAACCTCGCAGGACACTGTGAGGGTCATCGAACGGCTGATG
CTCTTCGTCGGCAGCTCATCAAGCACTGAAGGCCAGCAACTGGCCGCCGAGGCAAATGCGGTGCTCGGCATGAGCGTGGGA
GGCCTGCGCATGA

Protein sequence: (SEQ ID NO: 213)

MNLQNRNLLLSLIAETQFDAYVQGYMAKAGAAAGASENLQIEAEGAAMLQGLVAPVRAQQRACGQSLQNALQLIAHDL
QTKSQLAIAANASSIQVIQRDMNRAIWNIATAIDHLAEFAQPSQDTRVRIERLMLFVGSSSSTEGQQLAAEANAVLGMSV
GLA

RL114

DNA sequence: (SEQ ID NO: 93)

CTGAACAAGTTCGGCAGCGCCGCCGACCTTCGGAGCCAGCAGGCCAAATTGACCGGCGCTACGCGAGAAATACGCAAGCTG
ACTGGTGGCGGTATCGACCTGTTTCGGGAAGCTGGGTGCTACTTGAGCTTCGAACAAAAGCAGCTCCTACAAGACGCAGCG
CGCTTGCTCGACTCGGTGAACAAGCAGATCGAGCATGCGAAGGAAAAGCGTGATCGCTACGAGAAAAAGCCAGAAGCGG
CGCGAGCTACGTGAGCGCTGGCCAAGCAACTGGTGCCTCGAACTACCCGCTTCGGGAAATACGCTCGAAGATCGGCTG
GAAATCCTGCAGATCGCGTTGATCTACAACCGGGCCAGGGTGTTCGATCACCTGTACTCCACGCACCGCTCCACTCAAAA
CTCAAACGCTGGCTGGAGCGTCAAAGCAGCTCATCGGATGGCGCAGTGAAGCCGAGTATTTTCGCTAGTCAGGTGGGGAGC
CTGCGATGTGACTTCATTAGCCATCTGACTAACGAAATCGCGTACGACGATGGCAGTGAAGTCAGGAGCGCCTGCGCGTC
ATCAAGCAGAAGGTCGCTGACTGCACCGCACAGATCGCTCTGACCAGCGAGGAGCAGGAAACCCCTTCGCTCTGGACAGAC
GCTCTGCAATCGGCTCCGGAGGGCCTCATATGA

Protein sequence: (SEQ ID NO: 214)

LNKFGSAADLRSQAKLTGATREIRKLTGGGIDLFGKLGCYLSFEQKQLLQDAARLLDSVNKQIEHAKEKRDRYEKKAKKR
RELRLAKQLVASNYPLPGNTLEDRLEILQIALIYNRVRFDHLYSTHQLHSLKRWLERPKQLIGWRSEAEYFASQVGS
LRCDFISHLTNEIAYDDGSEVEERLRVIKQKVADCTAQIALTSEEQETLRLWTDALQSAPEGLI

67/118
FIG. 30V

RL115

DNA sequence: (SEQ ID NO: 94)

ATGAATGCGAAAGCGACTTCGGTTGTATCCACCAAGGGTGGTGTAGGAAAATCCACCACCGCCGCCAACCTCGGTGCATTTTG
CGCCGATGCAGGCATACGAACCTCCTCATCGATCTGGACCCCGTCCAGCCCTCCCTATCCTCGTACTACGAGCTGCCGGAAG
TTGCCCAGGGCGGCATTTACGACCTGCTCGCCGCCAACATAACGGACCCGGCGAGGATCATCTCCAGGACGATTATCCCAAT
CTGGACGTCGTGATTTCCAACGACCAGAACCAATCAGCTCAACAACCTACTGCTCCAGGCGCCCGATGGCCGGCTACGCCTGGC
GAACCTGATGCCCGCTCTGAAAGAAGGCTACGACCTGGTGTCTGATCGACACCCAGGGTGCGCGCTCAGCTTTGCTCGAAATGG
TTGTGCTTGATCGGACCTGGTTGTTTCCCCCTCCAACCCAACATGCTTACCGCCCGTGAGTTCAACCGCGGCACCATGCAA
ATGCTCGACGGCCTACGCCCTATGAGCGTCTCGGCATGCGGATCCCCAATGTTTCTGATCGTCATCAACTGCCTGGACCAGAC
CAATGACTCCCGGGCAATTCACGAGAATGTGCGTGCCATCTTCGATGAGCATCAGGACATTTCTGTGCTCGAAACGACTGTCC
CGGATGCCGTCGTGTTTCGCAACGCAGCATCGCGCGGGCTACCAGCGCACCGCCTCGAAACGCGGCAACCTCCAATCGCACA
TCAGCGCCCGCGCTGGAAATCATTGAAACCTGGCCATCGAGGTCTTTCCCGAGTGGACTGACCGCTTCCTGGCGCTGACGCC
GGGAGGCGGTTGCAGCACTGGTCAAGGGAGGGCGCTGACATGGCGAAGACTCCTATCACCCAAGCCCGCGACGTGACGCGGA
ACTTGTGCTGGAAGTGA

Protein sequence: (SEQ ID NO: 215)

MNAKATSVVSTKGGVGKSTTAANLGAFCADAGIRTLIDLDPVQPSLSSYYELPEVAQGGIYDLLAANITDPARIISRTIIPN
LDVVISNDQNNQLNNLLQAPDGRRLRLANLMPALKEGYDLVLIDTQGARSALLEMVVLASDLVVSPLQPNMLTAREFNRTMQ
MLDGLRPYERLGMRIPNVQIVINCLDQTNDRAIHENVRAIFDEHQDISVLETTVPDAVVFRNAASRGLPAHRLETRQPSNRT
SAPALEIIRNLAIEVFPEWTDRLALTPGGGCSTGQGRALTWRRLSPKPATSTRNLCWN

68/118
FIG. 31A

RS01

DNA sequence: (SEQ ID NO: 95)

ATGGGGATCTACCGCCCGAAGCAGTCTCGCCTAGCGATACCGATACTGAGGGGCGGCTACCGGACGAAAGGTAGCTGCG
CCTCCCAGCAGTTTCGCTAGGCCTGTAGGAAAAATCTGGAATTACCGAGAGCGCCTGGATTCCAGCGCCGGCATGCTGGCA
GAGCCCCGCAATTTCAAGGCCGAAACCGCAGTACCCTCTGTAATCGCTGATTACGTCGAGGGCACATTGCTACGCCTGCA
GAATGGTTTCAGGGCCTGAAAAACAGAAAAGCCACCTAAATAGGCGGGCTATTCCATATTGACATCACGTCAATGCGGG
CC

RS02

DNA sequence: (SEQ ID NO: 96)

ATGACGCCGAGCAGCTACCGAGGAGTACATCTTCGCGCACGATCTCCGAGAAGCCAGCGCGAAGATCTACCGCGCCGC
GACCAAGGCGCTGCTCAAGCACTTCGGCCCTACGGCAACCGTACAGGAGGTGGACCACAGGTCTGTGCTGGGATGGCGGC
GCAAGGTCCTGGAACAAGGCCTGTGCAAGCGGAGCTGGAACACGTATTGCAATCATCTGCGAACGATCTGGGGCTATGCC
ATCGAGCAGAGCTGGTGACACACTCCCAAGTCAACCCGTTCAAGAAAGACCACCGTCATCCCCCAGGCGAGCAAGCAA
AACCGTCGAGCCGAAGCCATCCTGCGCGCCCGCAATTGGCTCAACATGCAGGTGCGCGCCGAGCGCTGCACTGGCGATC
GCGCACGCATCACTCCCGCCTGGTTCTGGCTTTGCACGTTTGAGGTCTTCTACTTCACCGGCATCCGGTTGAATGCGCTG
TTGTGCATCCGCAAGCGCGACATCGACTGGGAAAATCAACTGATCCTCATCCGCGGCGAGACAGAGAAAACCTATAAAGA
GTTTCGTAGTGCCAATAACGGAGGGGCTTGTGCCTCACCTATCGCGGCTCCTGCAGGAGGCCGATAGAGCCGGATTTCGCGG
ATGACGACCAGTTGTTCAACGTCAACCGGTTCTCACCGCACTACAAGAGCAAGGTGATGAACTCCGACCAGGTGCAAGCC
ATGTACCGGAAGTTGACCGAGAAGGTTGGGGTGCGGATGACTCCGCACCGTTTCCGGCACACCCTGGCCACCGACTTGAT
GAAGGCACCCGAGCGGAACATCCACCTCACGAAGTGCCTGCTCAACCACTCGAATATCCAGACCACCATGAGCTACATCG
AGGCCGACTACGACCACATGCGTGCCGTGCTGCATGCCAGAAGCCTGGCCCAAGGAGCGCTGGAGAACGTGAGGAAGGTG
GATTACAGCGGCTCCCCGCAAGCCTCTGCCAAACCGAAGCCATGCGGGCAACCTCTCGCTCGAATGGGTGAAGCGCCGCC
ACAGGAGGCTAGGACAGAACCTGCAGAACCAAGGGAGCACACACAGGGACAGGCATTACAGGGAGATGCAACCGCGTGGG
AAGAAGCGCTACCACAGCCACCTGACACCTTCGAGCAAAGCGTGCTGTTCACTCTGATGGCTCAACACCTATCGAACCGT
GCCGCCACGGCTCCGCGGCTTCCACCGCAACAAGCGGATCTGGAGGATGGGGATCTACCGCCCGAAGCAGTCTCGCCTA
G

Protein sequence: (SEQ ID NO: 216)

MTPQQLTEEYIFAHLREASAKIYRAATKALLKHFGPTATVQEVDRHSVLGWRRKVLEQGLSKRSWNTYSNHLRTIWGYA
IEHELVTHSQVNPFRKTTVIPRRASKTVAAEAILRARNWLNMQVGAERCTGDRARITPAWFWLCTFEVYFTGIRLNL
LCIRKRDIDWENQLILIRGETEKTHKEFVVPITEGLVPHLSRLLQEADRAGFADDDQLFNVNRFSPHYKSKVMNSDQVEA
MYRKLTEKVGVRMTPHRFRHTLATDLMKAPERNIHLTKLLNHSNIQTMTSYIEADYDHMRAVLHARSLAQGALENVRKV
DYSGPSQASAKPKPCGQPLARMGEAPPQEARTPAEPREHTPGTGIQGDATAWEEALPQPPDTFEQSVLFTLMAQHLSNR
AATASAASTATSGSGGWGSTARSSLA.

RS03

DNA sequence: (SEQ ID NO: 97)

ATGAAATCTGGTATCGCGACCCGTCGCCTGTTTCATCAACGACACCAAGGCTTTGGTGCATACCGTCGACGGGACCGCCAT
GCTGGTCACGCCAGGAATCTTCAAGCGTTATGTCCAGGAGCATCCGGAGGTTGAAAAGCTGGCCCAGGCCAAGGAGACCG
CCGGCTGGAAGCTGGTGCAGCGCGCGTTCGAGAAACAGGGTCTTACCGAAAGACCAGTAAGAACCTGAATATCTGGACC
ATCAAGGTTTCTGGTCCTCGCAAGACGAAAGAGCTCAAGGCCTACCTGCTCCAGGATCCCAAATTGCTGTTCCTGTGCA
GCCTCTGGACAACCAAGCCTCACGGTCATCACCGATGCCGAAGGAGGTGTGGAATGA

Protein sequence: (SEQ ID NO: 217)

MKSGIATRRRLFINDTKALVHTVDGTAMLVTPGIFKRYVQEHPVEVEKLAQAKETAGWKLVQRAFEKQGLHRKTSKNLNIWT
IKVSGPRKTRELKAYLLQDPKLLFPVQPLDNPSTLTVITDAEGGVE

69/118
FIG. 31B

RS04

DNA sequence: (SEQ ID NO: 98)

ATAGACCAGTTGAGTGAGCAGGAGTCGGTGGAAGTGGTCTGCTCAGCTTTCGATGTGGCGCGGTCTTGCTACTACGTCCA
CCGTCTTCGACGCGCGGTGTGCGATGCTCGCCGCGTGGCGCTACGCAGCCAAGTCAACCAGTTGTTTCAGCCAGAGTCGGG
GCTCGGCCGCGCAGCCGAGCATTCTGGGCATGCTGCGCGAAGAGGGCGTGACCATCGGCCGTTTCCGAGTGCGTCCGGTTG
ATGCGTGAGTGGGCTGGTCAGCAAGCAACCGGGCTCGCACGCCTACAAACAGGCCACGGTTGAGCGGCCGATATCCC
GAATCGGCTGAACCGCAATTTCGCGACCGAGCATCCCATACAGGTGTGGTGTGGCGACATCACCTACGTCTGGGCGCAAG
GCCGTTGGCACTACCTGGCCGCGGTGCTGGATCTGCTGATCGGCTGGGCGTTCTCGGCCAAGCCGGATGCCGAACTGGTG
ATCAAGGCCCTGGACATGGCCTACGAACAGCGCGGCAGGCCACAGCAGGTGCTGTTCCATTTCAGACCAGGGCAGCCAGTA
CGCCAGCCGCTGTTTCGGCAACGGCTCTGGCGCTATCGGATGCAGCAGAGCATGAGCCGTCGGGGGAATTGCTGGGATA
ACTCGCCGATGGAGCGCCTGTTCCGCAGTCTGAAGTCGGAGTGGGTCCCGTCAACGGGTTACCTGACGGCGCAGGAGGCC
CAACGGGACATCAGTCATTACTTGATGCACCGCTACAACCTGGATCAGGCCGATCAATTCAACGACGGGTACCACCTGC
GGTGGCCGAAGAAAACTCAACCCACTGTCCGGGATGGGTTGA

Protein sequence: (SEQ ID NO: 218)

IDQLSEQESVEVVCSAFDVARSCYYVHRLRRRRVDARRVALRSQVNQLFSQSRGSAGSRILGMLREEGVITIGRFRVRL
MRELGLVSKQPGSHAYKQATVERPDI PNRLNREFATEHP IQVWCGDITYVWAQGRWHYLA AVL DLLIGWAFSAKPD AELV
IKALDMAYEQRGRPQQVLFHSDQGSQYASRLFRQLRWYRMQQSMSRRGNCWDNSP MERLFRSLKSEWVPSTGYLTAQEA
QRDISHYLMHRYNWIRPHQFNDGLPPAVAAEEKLNPLSGMG

RS05

DNA sequence: (SEQ ID NO: 282)

ATGAGCAAGCAACGACGTACGTTTTCCGCCGAGTTCAAACGAGAGGCCGCGGCCCTGGTGTGGACCAAGGCTACAGCCA
TATCGACGCCCTGCCGTTTCGCTGGGGGTGGTGGATTCCGCCCTTCGCGCGTTGGGTGAAGCAGCTCGAGGCGGAGCGCCAGG
GTGTGACCCCGAAGAGCAAGGCGTTGACGCCCTGAGCAGCAAAAGATCCAGGAGCTGGAAGCCCGGATCAACCGATTGGAG
CGGGAGAAAGCGATATTAAAAAAGGCTACCGCTCTCTTGATGTCCGACGAACTCGATCGTACGCGCTGA

Protein sequence: (SEQ ID NO: 219)

MSKQRRTFSAEFKREAAALVLDQGYSHIDACRSLGVVDSALRRWVKQLEAERQGVTPKSKALTPEQQKIQELEARINRLE
REKAILKKATALLMSDELDRTR

RS06

DNA sequence: (SEQ ID NO: 99)

ATGTTGTATTTTCTTGCAAGATGAGATGGGTGGTGGGTCCGATATAGGTACTTCTCTCTATTTTCTTTAATTGCTCT
CATCTATGGGTGTGTCCGTGGTGGAGGTGGATCGGATGAGATTGGGCAGCACTGCTTTGAGAGAGAGCAAAAGCTTTCCG
GAGTTAATGATAATGAAGAGGGGAGTGTGAGGTTGAATCGGCTGAACTGCGATCCAATTGAAGGTCGTGTTCTTGAATCA
GAGAAGCTGATAAGAAAGCCGCCCAATGAGCTGGGTATTCACTGA

Protein sequence: (SEQ ID NO: 220)

MLYFSCSMKMGWVG YRYFSLFSLIALIYGC VGGGGGSDEIGQHCFEREQKLSGVNDNEEGSVRLNRLNCDPIEGRVLES
EKLIRKPPNELGIH

70/118
FIG. 31C

RS07

DNA sequence: (SEQ ID NO: 100)

ATGAAAAAATCACTTGTATGTCTGGCTGTGCTTTTGGTGGCTAGCAATTTGCGGTGTGCTGATGAGGGCTCAAATGA
TGGAAGTGAGATATGTCGGGCGCAGGGTGGAGTTGAAATAACAAGTCTGGGGGAAGTCTCAAAGGGTGTGGATGTTG
AAGATGTTGTAGTTTGTTCGATTCTTCCAAGTAATATGAAGTCGAGTCAAAGAGCGCCTACACTCCCTCCTCTGCAA
AGGATGATCATTTTCGGCAATGCCTTCACCAGGAACGGTCACTGTTTCTGCCAGCGGAGATAGGAAATTTACAACATC
TTGCCGGGGCAAATCTTTATGCTCCACGTTATGCCAATTTCTATCCAGACGGTGTAGCAGGGGAACATCAGATCTAC
GATGTGTTGGTTACAATACACCCGGAATTCATCTCAAGGGTGTAAATGTGTGCATGGGACGGCCCCGACCGACATTCAA
TTGGGTGTTGAGCCATATGGCGGATCTGTTGTTGTAACTACAGTTGCACTGCATTCAAACAACGATTCCAGTGAT
AATGAGCTACAGTTATCGTGATGGGCGGGCAGTGATGGCGAGGTCCAGAATGTGTGTCAGGAATAATAAATGTGGTTT
TGAACTAA

Protein sequence: (SEQ ID NO: 221)

MKKSLVMSAVLLVASNFACADEGSNDGSEICRAQGGVEITSLGEVSKGVDVEDVVVCSILPSNMKSSQRAPTLPLQ
RMIIISAMPSPGTVTVSASGDRKFTTSCRANLYAPRYANFYPDGVSRGTSDLRCVGYNTPGNSSQGCNVSWDGPDIQ
LGVEPYGGSVVVNYSTAFKTTIPVIMSYSYRDGRAVYGEVQNVSGIINVVLN

RS08

DNA sequence: (SEQ ID NO: 101)

ATGCTTATTAAAATTCTTCGAATTATATTCTTGTTGCCTATAGTTGGTTTGGCACAGCAGGCTGCTGCCTCCCCGCC
CGCAGAGTCACACTCGGAACAATCTGAATCTTCGTGTATCGATGTCCAAGTCAATGGAGCACGTAGCCTGTCTTATA
ACTGCATGGCTCAGCAAATGACTCCACCCAAAGAGGATCCTCGGCGTCGGAACCCCTACCTTGAATCCACATTAGCG
TCTGAACGCGCCACTCGCCTGCCACCCACACAGACAGGACTTTTTACCAGCCTTCATCAACGTGCCATATCGAACTC
GAAAGACTAG

Protein sequence: (SEQ ID NO: 222)

MLIKILRIIFLLPIVGLAQQAASPPAESHSEQSESSCIDVQVNGARSLSYNMAQQMTPPKEDPRRRNPTLNSTLA
SERATRLPPTQTGLFTSLHQRAISNSKD

RS09

DNA sequence: (SEQ ID NO: 102)

GTGAGTAGTACTAAGAGTAAGCCGATAGCCAGGGGGCGTGGTGGCCCATTTGGGGAAGTGATGAAGAGGTGCGGGCT
TGTACCGGTTTCGAGGAAGGAATAGACAGCAGACAGGATCGCTTGCGATGGGGCAGCAGGAAACCATCAGCCCGTCCG
TATCCAGAACTGCTGCTTGCAGCGTTAGGGGTGACTCCCTCATGCCCTAG

Protein sequence: (SEQ ID NO: 223)

VSSTKSKPIARGRGPFGEVMKRCGLVPVRGRNRQQTGSLAMGQQETISPSVSRTAACSVRGDSLMP

RS10

DNA sequence: (SEQ ID NO: 103)

ATGGAACGCTTGCTCGAGAGCATTTACATCAATGCCCGGCCGGCGATGGAGTTGAGGCTTAGCCTCACCAGCTCCGG
CCGCAAGAGAATGGTAAAGATTGTGGATGGGGAGGAGGTCGAGGTTCTGCCAGGTGAAGTGCAGGGCATCTGGAGG
CCCAAAGAGGGATGTTGAATCCTCGCCGACTTCTTAGCCAAGAGTCTCGTGGCGCGACGCTAG

Protein sequence: (SEQ ID NO: 224)

MERLLESIYINARPAMELRLSLTSSGRKRMVKIVDGEEVEVLPGEVQGILEAQKRDVGILADFLAKSLVARR

71/118
FIG. 31D

RS011

DNA sequence: (SEQ ID NO: 104)

ATGGAATGCCACGTTCTCGTCCCGCCACGAGCAGAGATGCAGCAGCGATAAGCTGCGTAGTTATAGCCGCCCTGCGTGA
GTCAAATTCACAGGACTATCCGCCTGATGTGATCGCTCAGGTTGAGCAGAGCTTTTCTCCTGAAGCCATCACCACAC
AGCTTACGAAGCGTAGGGTCTTCTGAGCCTTATTGGGCGAAAACATTATTGGCACTGCCGGTCTCGACGGTGACGTC
GTCAGAAGTGTTCGTTGACCCAGCTCACCAGAAAGGCGGTATCGGGCGGCATTGTGATGGATGTCATTCATACAAC
TGCTGCCAGCGCGGGAGTTGGAGCTGTACGTGTGCCATCGTCGATTACAGCTGAAAAGGTTTATACCGCATTGGGTT
ATCAGAAAATCCGCGACGAGTTTCATGGGGCGGAGCGCACCATCGTTATGGAGAAGCGGCTGTAG

Protein sequence: (SEQ ID NO: 225)

MECHVRPATSRDAAAISCVVIAALRESNSQDYPPDVIAQVEQSFSP EAITTQLTKRRVVFVALLGENIIGTAGLDGDV
VRSVFDPAHQKGGIGRHLMDVIHTTAASAGVGA VRVPSSITAERFYTALGYQKIRDEFHGAERTIVMEKRL

RS12

DNA sequence: (SEQ ID NO: 105)

TTGTGGTTGACCTGCACGCCACAGCAGGATGTGCAGGCGGCGTTAGCTACAGCGTCGATACTCCTGGGCCAGTTCCA
CCAGTTGGGCGTGACGCTCGGTTCGGTACACTAGCCTCGACCCGCTTGAGGAAGTCGAGAAGAAGCCTTCTGCACTGC
CGTCTCCTGCTTGAAAAACGATTCTACTAAGTTCAGCGTGGTACTGAAATCGGGGGGCAGGTCAATCGACAAAAGGT
ATCCCGACCGCAGGTTTGTGGCCACGTCATGGTGGCCAAGTTTGCCGATCACTTGCCGCTGTACCGGCAGGAGAA
AATCTTTGGCCGCGCCGGGCTGGCAATTGCTCGCTCGACCTGGCGCAGTGGGTCCGACAAACCGGCGTGCGGCTTC
AGCCACTGGTCGATGCACTGCGTGAAGCCGTGCTGAACCAGGGCGTGATCCACGCTGATGAAACACCGGTGCAAATG
CTTGCGCCAGGCGAGAAGAAAACCCACCGGCCCTATGTCTGGGCGTACAGCACGACGCCGTTTTCAGGGCTCAAAGC
GGTGGTTTACGACTTACGCCAAGCCGTGCTGGCGAACATGCGCGCAACTTCCTGGGTGACTGGAACGGCAAGCTGG
TCTGCGACGACTTCGCTGGCTACAAAGCCGGTTTCGAACAAGGCATCACTGAAATCGGCTGCATGGCCACGCCCGG
CGCAAGTTCTTTGATTTGCACGTGGCGAACAAAAGTCAGCTGGCTGAACAGGCCCTGCACTCGATCAGCGGCTTGTA
CGAGGTGCAACGTCAGGCGCGGGACATGAGTGATGAAGAGCGCTGGCGAATACGACAAGAATTGGCGGTGCCGATCC
TCAAAAACTGCATGACTGGATGTTGGCTCAGCGAGACCTGGTGCCCAATGGATCAGCCACGGCCAAAGCCCTCGAT
TACAGCCTGAAACGCTGGGTAGCGCTGACGCGCTACCTGGACGATGGGGCTGTGCCCATCGATAACAATCAGGTCGA
GAACCAAATACGGCATCGGCGCTCGGGCGTTTCAACTGGCTGTTTGCCGGGTGCTGCGCAGTGGTAAACGGGCGG
CTGCAATCATGAGCCTGATCTAG

Protein sequence: (SEQ ID NO: 226)

LWLTCTPQQDVQAALATASILLGQFHQLGVQLGRYTSLDPLEEVEKNASALPSPAWKTDSTKFSVVLKSGGRSIDKG
IPTAGLLAHVMVAKFADHFLPLYRQEKIFGRAGLAIARSTLAQWVGQTGVRLQPLVDALREAVLNQGVIIHADETPVQM
LAPGEKKTHRAYVWAYSTTPFSGLKAVVYDFSPSRAGEHARNFLGDWNGKLVCDDFAGYKAGFEQGITEIGCMAHAR
RKFFDLHVANKSQLAEQALHSISGLYEVEERQARDMSDEERWIRIQELAVPILKKLHDWMLAQRDLVPNGSATAKALD
YSLKRWVALTRYLDDGAVPIDNNQVENQIRPWALGRSNWLFAGSLRSGKRAAAIMSLI

RS13

DNA sequence: (SEQ ID NO: 106)

ATGGTGAGGCGGCGGAGGGTCGCGGTGGCGCGCAATGCCTGAGCCTGTGAGCGCACCGAACCAGGTCTTGTCGAT
GGATTTCTGCTTCGACGCGCTCAGCACTGGGCGACGGATCAAATGCCTGACGGTGGTTCGATGACTTCACCAAGGTGT
CGGTTCGACATCTTGGTGAGTACGGTATCAGCGGTTTTCTGTGTACGCGGGCGCTGGACGAGATGGCGCGGTTTCGT
GGTACCCGCGAGCGATCCGACCGACGAGGCCCCGAGTTACCGGCAAGGCGCTTGATCAGTGGGCTGTGTCAGCG
TGACATCAAGTTGAAGCTGATTCAGCCTGGCCAGCCACGACGAGCGCCTTCATCGAGTCATTCAACGGCAAGTTCC
GGGGCGAATGCCTCAATGAGCACTGCTCGCTGGTTCGAAGCCAGAATCCGTATCGCGGCTTGGCGGGATTACAACGAG
CACCGACCACACAGCGCCATTGGCAATCTCTCCCCGGCAGAGCTTGCTGCGAAGTGGCGAACCAACCAGCAGCAGCT
GAAGCGGGAAAAGTTGATATCAACCCCATAG

Protein sequence: (SEQ ID NO: 227)

MVRRRRVAVARECLSLSSAPNQVLSMDFVFDALSTGRRIKCLTVVDDFTKVSVDILVEYGISGFRVTRALDEMARFR
GYPQAIRTDQGPEFTGKALDQWACQDIKLIKLIQPGQPTQSAFIESFNGKFRGECLNEHCSLVEARIRIAAWRDYNE
HRPHSAIGNLSPAELAAKWRTNQQLKREKLSTP

72/118
FIG. 31E

RS14

DNA sequence: (SEQ ID NO: 107)

ATGCATATCCAATCGTTGGGGGCTACTGCCTCCTCGCTGAATCAGGAGCCTGTGCGAAACCCCGTCGCAGGCAGCGCATAA
 GTCCGCCAGCTTGCCTCAGGAACCTTCAGGGCAAGGTCTCGGGGTTGCCCTAAAGAGCACGCCGGGAATACTTTCCGGGA
 AGTTGCCGGAAGCGTTAGCGACGTGCGTTTCAGCAGTCCCCAAGGGCAAGGGAGTCCCGTACTCTGACTGACTCGGCA
 GGGCCGCGGCAGATCACTCTGCGCCAGTTTGTAGAACGGAGTACCGAGCTACAGCTCAGTCGGCCACCATTGACCAGTCT
 GGTCTTAAGCGCGGTGGTGCCAAAGGTGCGGCATACCCGGGAGCAATGCTGGCGCTAGAAGAGAAAGGCATGCTCGATG
 GCATCCGCAGCATGTCCGGTTCGTCCGCTGGCGGCATCACCGCCGCCCTTTTGGCCTCAGGTATGAGCCCGCGCGGCTTC
 AAGACCTTTCCGACAAGATGGATCTTATTTTCGTGCTCGACAGCTCGAACAAGAAGCTGAAGCTGTTTCAACACATTAG
 CAGCGAGATCGGCGCATCGCTGAAAAAGGGCTTGGGCAACAAGATCGGCGGCTTCTCTGAGTTGCTGCTCAATGTACTCC
 CACGCATAGATTTCGCGGGCTGAGCCCCTAGAACGCCCTATTGCGCGACGAGACACGCAAGGCCGTGCTCGGACAGATCGCT
 ACGCATCCAGAGTTGCACGCCAGCCGACCGTTGCCGCCATCGCCAGCAGATTGCAGTCCGGCTCCGGAGTCACCTTTGG
 CGATCTAGATCGGTTGAGTGCTTACATTTCCCAGATTAAAGACGCTGAACATCACAGGTACGGCCATGTTTCGAGGGGCGTC
 CGCAATTAGTGTTCAATGCCAGCCACACACCGGATCTGGAGGTGCGCCAGGCGGCACATATCTCCGGTTCCTTCCCA
 GGAGTGTTCCAGAAGGTGAGTTCAGTTCAGCGCTACCGGCGGCGTAGAGTGGACAGAATTCAGGATGGCGGGCT
 GATGATTAACGTGCGCGTCCCTGAGATGATCGACAAGAATTTTGACAGCGGGCCACTGCGGCGCAACGACAACCTGATCC
 TTGAGTTCGAGGGCGAAGCTGGGGAGGTAGCGCCGACCGAGGTACTAGGGGCGGCGCGCTCAAGGGCTGGGTGCTCGGG
 GTGCCCTGCCCTGCAGGCGCGGAAATGCTGCAGCTCGAGGGCTTGAGGAATTGCGCGAGCAAACCGTTGTGGTGCCGTT
 GAAGAGCGAGCGCGGTGATTTTCAGTGGCATGCTCGGTGGCACCTTGAACCTTCACCATGCCGGACGAGATCAAGGCGCATC
 TTCAGGAGCGCTCCAGGAGCGAGTCCGTGAACATCTGGAGAAACGTCTTCAGGCTTCAGAGCGTCATACCTTCGCTTCT
 CTCGACGAGGCGCTGCTGGCACTTGATGACAGTATGCTCACCAGTGTGTGCTCAACAGAACCCGGAGATCACAGACGGGGC
 GGTGGCTTTTTCGCCAGAAGGCGCGGATGCGTTACCGAGCTGACTGTGCTATCGTTAGCGCCAATGGCTTGGCGGGTA
 GGCTCAAGTTGACGAGGCTATGCGCTCCGCTCTTCAGCGACTCGATGCGCTGGCAGATACTCCGGAACGCCTAGCATGG
 TTGGCAGCTGAGTTGAACCATGCTGATAACGTTGATCATCAGCAGTTACTCGATGCCATGCGCGGGCAGACGGTGACGTC
 GCCGGTGCTCGCGCTGCGTTAGCAGAGGCGCAGCGCCGCAAAGTGCGGTTATTGCCGAGAACATTCGTAAGGAAGTTA
 TCTTCCCCCTCTCTGATCGCCCTGGCCAGCCGATTCCAACGTAGCTCTGTTACGTGCGGCGGAGGAGCAGCTACGGCAT
 GCCACAGTCCGCGGAAATCAATCAAGCGCTGAACGATATCGTCGACAATACTCGGCACGAGGCTTCTGCGTTTCGG
 CAAACCTTGAGTTCGACTACCGTTGAGATGGCTAAGGCTTGGCGGAATAAGGAGTTACATGATT

Protein sequence (SEQ ID NO:228)

MHIQSLGATASSLNQEPVETPSQAAHKSASLRQEPSPGQGLGVALKSTPGILSGKLPESVSDVRFSSPQGQGESRTLTD
 SA GPRQITLRQFENGVTSLQSRPPLTSLVLSGGGAKGAAYPGAMLALBEEKGMLDGI RSMSSAGGIT AALLASGMSPAAF
 KTLSDKMDLISLLDSSNKKLKFQHI SSEIGASLKKGLGNKIGFSELLNLNVLPRIDSRAEPLERLLRDETRKAVLQIOIA
 THPEVARQPTVAAIASRLQSGSGVTFGDLDRLSAYIPQIKTLNITGTAMFEGRPQLVVFNASHTPDLEVAQAAHISGSFP
 GVFKVSLSDQPYQAGVEWTEFQDGGVMINVPVPEMIDKNFDSGLPRNRNDNLILEFEAGEVAPDRGTRGGALKGWVVG
 VPALQAREMLQLEGLEELREQTVVPLKSERGDFSGMLGGTLNFTMPDEIKAHLQERLQERVGEHLEKRLQASERHTFAS
 LDEALLALDDSMILTSAQQNPEITDGAFAFRQKARDAFTELTVAIVSANGLAGRLKLDEAMRSALQRLDALADTPERLAW
 LAELNHADNVDHQQLLDAMRGQTVQSPVLAALAEARRKVAVIAENIRKEVIFPSLYRPGQPDNSVALLRRAEEQLRH
 ATSPAELNQLNDIVDYSARGFLRFGLKPLSSTTVEMAKAWRNKEFT

RS15

DNA sequence: (SEQ ID NO: 108)

ATGATTGATACATGGCTGGCACAGTGGGGCTTGAGACTTCCCTCGAGCAACGATGCCACGTTGCGGCTGCAACCGGCAGA
 GGGACCGGAACGTGTTATGGAGCGCCTCGAGGGCGGTTGGCTTTTCGTGCTGAGTTGGGACTTGTGCCTTCAGGGTTAC
 CGCTGGGTGTGATCTTGCAATTGTTACAAGTGAACCTCTCCATTCTCATCCTTGGCACCGGTGAACTTGGCGGGACGAT
 GCCGGTAGACTTGTGCTCTGGGCTGAGGCACGTGATGGCGTTGACGATGTGGATGCACTGAACCGCTTGACGATAGGCT
 GCGGGAAGGACATTACGATTAGTGCCATTGCTAGAGCCCACGGTGAGTTGGTTCCAGCTCAGATACAAACCAGCGCGT
 TAGTGTTCGTTTGA

Protein sequence: (SEQ ID NO: 229)

MIDTWLAQWGLRLPSSNDATLRLQPAEGPELVMERLEGGWLFVVELGLVPSGLPLGVILQLLQVNSPSSSLAPVKLAADD
 AGRVLVWAEARDGVDDVDALNRLHDLRLREGHSRLVPLLEPTGELVPAQIQTSALVFV

73/118
FIG. 32

Big Island: Overall Nucleotide Homology

Total 84830bp

Bp#	Species, strain, gene name	Accession #	Evalue/ %identity
-----	----------------------------	-------------	-------------------

23559-25465:	X. axonopodis pv. Citri strain 306	AE011864	83%
--------------	------------------------------------	----------	-----

Only stretches within the 20437-25465bp are homologous to X. axonopodis pv. Citri strain 306.

A total of 1060bp, not contiguous, from this region are homologous to X. axonopodis.

33872-38412:	P. aeruginosa, PA14, pvrR	AF482691	0.0; 99%
--------------	---------------------------	----------	----------

40989-46535:	P. aeruginosa, PA01, PA2128-2132	AE004640	80%
--------------	-------------------------------------	----------	-----

Only stretches within the 40989-46535bp region are homologous to PA01. A total of 2406bp, not contiguous, from this region are homologous to PA01.

48266-49533:	P. putida, plasmid pWWO	AJ344068	96%
--------------	-------------------------	----------	-----

Only stretches within the 48266-49533bp are homologous to P. putida, plasmid pWWO. A total of 780bp, not contiguous, from this region are homologous to P. putida.

56824-58706:	P. syringae pv. maculicola, plasmid pFKN	AF359557	83%
--------------	---	----------	-----

Only stretches within the 56824-58706bp are homologous to P. syringae pv. maculicola, plasmid pFKN. A total of 1882bp, not contiguous, from this region are homologous to P. syringae.

64748-64942:	P. aeruginosa, PA103, exoU,	U97065	1E-85/96%
82447-85179	P. aeruginosa, PA01, PA0984-0985	AE004531	0.0; 97%
85334-855542:			3E-80/94%
93200-93317:	P. aeruginosa, PA158	X73064	7E-50/98%
108075-108610:	P. aeruginosa, SG17M, plasmid pKLC102	AF285416	0.0/91%
100119-101054:	P. aeruginosa, PA01, PA3849	AE4802	0.0/98%

FIG. 33

Small Island: Overall Nucleotide Homology

Total 10848bp

Bp#	(Species, strain, gene name	Accession #)	Evalue/ %identity
877-1632:	<i>P. aeruginosa</i> , PA01, PA0977	AE004531	0.0/94%
877-2603:	<i>P. aeruginosa</i> , SG17M plasmid pKLC102	AF285416	0.0/92%
2978-6471:	<i>P. aeruginosa</i> , PA01, PA0978-81	AE004531	0.0/99%
7035-7999:	<i>P. syringae</i> pv. <i>maculicola</i> , plasmid pFKN	AF359557	83%

Only stretches within the 7035-7999bp are homologous to *P. syringae* pv. *maculicola*, plasmid pFKN. A total of 534bp, not contiguous, from this region are homologous to *P. syringae*.

7999-8284:	<i>P. aeruginosa</i> , PA103, <i>exoU</i>	AF27291	E-136/96%
8000-8080:	<i>P. aeruginosa</i> , PA01 intragenic region	AE004531	E-18/91%
8120-8259:	<i>P. aeruginosa</i> , PA01 intragenic region	AE004531	E-24/85%
8272-8860:	<i>P. aeruginosa</i> , PA01 intragenic region	AE004531	E-176/88%
8470-11724:	<i>P. aeruginosa</i> , PA103, <i>exoU</i> ,	AF27291	0.0/99%

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND
PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September 12, 2003 Serial No.: Not Yet Assigned

Page 74 of 118

Customer No.: 21559

FIG. 34A

ORF ID	Strand	Left end	Right end	ORF length (aa)	G+C content (%)	Location prediction	Gene name	Gene function	Protein with the highest identity (gene name / species strain)	E-value (% identity)	Proteins with lesser identity Cut off 30%	GenBank accession no.
	+	725	800		51				tRNA-Lys / <i>P. aeruginosa</i> PAO1			AE004888
RL001	-	878	1,202	pseudogene	56			Hypothetical protein	PA0977 / <i>P. aeruginosa</i> PAO1	3E-45 (91)		AAG04366
RL002	-	1,169	2,452	427	60	cytoplasm	xerC	Integrase	Integrase / <i>P. aeruginosa</i> SG17M	0.0 (91)	STY4666	AAG02084
RL003	-	2,449	4,368	639	58	cytoplasm		Hypothetical protein	PSPT00831 / <i>P. syringae</i> pv. tomato DC3000	E-180 (52)	XAC2196, XCC312, STY4665	AA054371
RL004	+	4,903	5,436	177	56	cytoplasm	dod_2	Deoxycytidine deaminase	MK1566 / <i>Methanopyrus kandleri</i> AV19	1E-13 (33)		AAM02779
RL005	+	5,436	6,146	236	51	inner membrane		Hypothetical protein	No significant similarity			
RL006	+	6,143	6,682	179	45	cytoplasm	dtd	Deoxycytidine triphosphate deaminase	VNG0245G / <i>Halobacterium</i> sp. NRC-1	1E-07 (32)		AAG18843
RL007	-	6,887	7,402	171	48	outer membrane and periplasm		Hypothetical protein	No significant similarity			
RL008	-	7,829	9,760	643	60	cytoplasm		Hypothetical protein	Protein fused from putative helicase (<i>Methanosarcina acetivorans</i> C2A) and hypothetical protein PA1935 (<i>P. aeruginosa</i> PAO1)	7E-09 (25) / 8E-16 (32)		AAM05538 / AAG05323
RL009	-	9,757	12,180	807	60	inner membrane		Hypothetical protein	PA1939 / <i>P. aeruginosa</i> PAO1	2E-69 (30)		AAG05327
RL010	-	12,358	12,660	100	54	inner membrane		Hypothetical protein	No significant similarity			
RL011	-	13,102	13,452	116	60	cytoplasm	parE	Plasmid stabilization protein parE	AGR_C_2415p / <i>Agrobacterium tumefaciens</i> C58	8E-22 (45)		AAK87104
RL012	-	13,456	13,728	90	62	cytoplasm		Putative transcription regulator	AGR_C_2413p / <i>A. tumefaciens</i> C58	2E-13 (41)	STY3093	AAK87103
RL013	+	13,847	14,191	114	48	inner membrane		Hypothetical protein	No significant similarity			
RL014	-	14,217	15,728	503	58	inner membrane		Hypothetical protein	XAC2186 / <i>X. axonopodis</i> pv. citr 306	E-175 (55)	XCC3117, STY4579	AAM37039
RL015	-	15,725	16,066	113	62	inner membrane		Hypothetical protein	No significant similarity			
RL016	-	16,066	17,448	460	64	outer membrane and periplasm		Hypothetical protein	PSP00848 / <i>P. syringae</i> pv. tomato DC3000	E-138 (57)	XCC3116, STY4577	AA054383

FIG. 34B

RL017	-	17,466	18,404	312	65	outer membrane and periplasm		Hypothetical protein	PSPO0849 / <i>P. syringae</i> pv. tomato DC3000	E-125 (72)	XAC2283, STY4576	AAO54384
RL018	-	18,404	18,835	143	62	outer membrane and periplasm		Hypothetical protein	SG52 / <i>P. aeruginosa</i> SG17M	3E-22 (49)	STY4575	AAM37135
RL019	+	19,044	19,262	72	55	cytoplasm		Hypothetical protein	No significant similarity		XAC2275	
RL020	-	19,259	19,918	219	62	outer membrane and periplasm	<i>dsbG</i>	Putative protein-disulfide isomerase	PA0982 / <i>P. aeruginosa</i> PAO1	1E-92 (93)		AAG04371
RL021	-	19,915	20,199	94	57	cytoplasm		Hypothetical protein	PSPT00858 / <i>P. syringae</i> pv. tomato DC3000	2E-23 (55)	STY4573, STY4572	AAO54393
RL022	-	20,196	23,138	980	64	cytoplasm		Hypothetical protein	PSPT00859 / <i>P. syringae</i> pv. tomato DC3000	0.0 (64)		AAO54394
RL023	-	23,138	23,581	147	64	inner membrane		Hypothetical protein	PSPT00860 / <i>P. syringae</i> pv. tomato DC3000	2E-48 (74)		AAO54395
RL024	-	23,559	25,064	501	63	outer membrane and periplasm		Hypothetical protein	PSPT00861 / <i>P. syringae</i> pv. tomato DC3000	E-148 (54)	STY4570, ORF109	AAO54396
RL025	-	25,048	25,932	294	66	outer membrane and periplasm		Hypothetical protein	PSPT00862 / <i>P. syringae</i> pv. tomato DC3000	1E-90 (62)	XAC2272, STY4569	AAO54397
RL026	-	25,929	26,588	219	60	inner membrane		Hypothetical protein	PSPT00863 / <i>P. syringae</i> pv. tomato DC3000	1E-71 (59)	STY4568	AAO54398
RL027	-	26,585	26,971	128	65	inner membrane		Hypothetical protein	XAC2271 / <i>X. axonopodis</i> pv. citri 306	4E-29 (50)		AAM37124
RL028	-	26,982	27,338	118	60	inner membrane		Hypothetical protein	ORF116 / <i>P. putida</i> (plasmid pWWO)	6E-23 (50)	XAC2270, STY4566	CAC86817
RL029	-	27,356	27,595	79	63	inner membrane		Hypothetical protein	C54 / <i>P. aeruginosa</i> C	1E-09 (38)	STY4565	AAN62148
RL030	-	27,592	27,951	119	66	inner membrane		Putative type III effector Hop protein	PSPT00869 / <i>P. syringae</i> pv. tomato DC3000	4E-25 (53)	XAC2270, STY4566	AAO54404
RL031	-	28,024	28,329	101	57	cytoplasm		Hypothetical protein	No significant similarity		PA4736, PA4737	
RL032	+	28,502	28,813	103	49	outer membrane and periplasm		Hypothetical protein	PA0713 / <i>P. aeruginosa</i> PAO1	9E-16 (43)		AAG04102
RL033	-	28,810	29,967	385	47	cytoplasm		Hypothetical protein	No significant similarity			
RL034	-	30,098	31,579	493	59	inner membrane		Putative DNA helicase	PSPT00879 / <i>P. syringae</i> pv. tomato DC3000	0.0 (65)		AAO54413

FIG. 34C

RL035	-	31,590	32,240	216	62	inner membrane		Hypothetical protein	PSPT0880 / <i>P. syringae</i> pv. tomato DC3000	2E-74 (63)	XAC2260, STY4563	AAO54414
RL036	+	32,572	35,367	931	60	inner membrane		sensor of two-component regulatory system	ORF1 / <i>P. aeruginosa</i> PA14 // RcsC / <i>S. enterica</i> subsp. <i>enterica</i> ser. Typhi CT18	0.0 (91) / 8E-98 (32)	VieS?	AAM15532 / CAD07502
RL037	+	35,364	36,563	399	59	cytoplasm	<i>pvrR</i>	regulator of two-component regulatory system; adhesion and antibiotic resistance	<i>PvrR</i> / <i>P. aeruginosa</i> PA14 // <i>VieA</i> / <i>X. campestris</i> pv. <i>campestris</i> ATCC 33913	0.0 (100) / 2E-55 (34)	<i>pvrR</i>	AAM15533 / AAM41975
RL038	+	36,644	39,898	1,084	63	inner membrane	<i>rscC</i>	sensor of two-component regulatory system	STM2271 / <i>S. typhimurium</i> LT2	4E-85 (32)		AAL21172
RL039	+	39,898	40,593	231	59	cytoplasm	<i>rscB</i>	regulator of two-component regulatory system	Z3476 / <i>E. coli</i> O157:H7 EDL933	8E-33 (39)		AAG57352
RL040	-	40,637	41,353	238	67	outer membrane and periplasm	<i>cupD5</i>	Probable pili assembly chaperone / adhesion and protein secretion	PA2132 / <i>P. aeruginosa</i> PAO1	4E-65 (62)		AAG05520
RL041	-	41,343	42,689	448	63	outer membrane and periplasm	<i>cupD4</i>	adhesion and protein secretion	PA2131 / <i>P. aeruginosa</i> PAO1	0.0 (70)	STY0370	AAG05519
RL042	-	42,686	45,298	870	66	outer membrane and periplasm	<i>cupD3</i>	Probable fimbrial biogenesis usher / adhesion and protein secretion	PA2130 / <i>P. aeruginosa</i> PAO1	0.0 (67)	STY0371	AAG05518
RL043	-	45,282	46,028	248	65	outer membrane and periplasm	<i>cupD2</i>	Probable pili assembly chaperone / adhesion and protein secretion	PA2129 / <i>P. aeruginosa</i> PAO1	1E-83 (64)	STY0372	AAG05517
RL044	-	46,117	46,665	182	65	outer membrane and periplasm	<i>cupD1</i>	Probable fimbrial precursor / adhesion and protein secretion	PA2128 / <i>P. aeruginosa</i> PAO1	6E-66 (72)	STY0373	AAG05516
RL045	-	46,836	47,101	pseudogene	55			Recombination	Transposase / <i>E. coli</i> (plasmid p1658/97)	2E-36 (82)		AAO49572
RL046	-	47,103	47,849	248	61	inner membrane		Hypothetical protein	PSPT0880 / <i>P. syringae</i> pv. tomato DC3000	1E-83 (59)	XAC2260, STY4563	AAO54414

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September 12, 2003 Serial No.: Not Yet Assigned

Page 77 of 118

Customer No.: 21559

FIG. 34D

RL047	-	47,849	50,083	744	65	inner membrane	Hypothetical protein	SG68 / <i>P. aeruginosa</i> SG17M	0.0 (66)	XAC2259, STY4562	AAN62290
RL048	-	50,087	50,344	85	59	cytoplasm	Hypothetical protein	No significant similarity			
RL049	-	50,353	50,853	166	66	outer membrane and periplasm	Hypothetical protein	SG69 / <i>P. aeruginosa</i> SG17M	7E-34 (51)	STY4560	AAN62291
RL050	-	50,850	51,431	193	64	outer membrane and periplasm	Hypothetical protein	SG70 / <i>P. aeruginosa</i> SG17M	5E-38 (46)	STY4559	AAN62292
RL051	-	51,416	52,171	251	65	outer membrane and periplasm	Hypothetical protein	C68 / <i>P. aeruginosa</i> C	5E-50 (52)	STY4558	AAN62162
RL052	-	52,182	52,871	229	64	inner membrane	Hypothetical protein	C69 / <i>P. aeruginosa</i> C	5E-05 (25)		AAN62163
RL053	-	53,019	53,235	pseudogene	61		Recombination	Transposase / <i>P. putida</i> (plasmid pWW53)	7E-09 (54)		BAB59051
RL054	+	53,482	54,513	343	47	cytoplasm	Hypothetical protein	PA2223 / <i>P. aeruginosa</i> PAO1	2E-56 (45)		AAG05611
RL055	+	54,524	55,189	221	50	cytoplasm	Hypothetical protein	PA2222 / <i>P. aeruginosa</i> PAO1	2E-56 (48)		AAG05610
RL056	+	55,272	55,946	224	49	cytoplasm	Hypothetical protein	PA2224 / <i>P. aeruginosa</i> PAO1	2E-06 (23)	XAC4135	AAG05612
RL057	+	56,030	56,275	81	59	cytoplasm	Hypothetical protein	No significant similarity			
RL058	-	56,348	56,707	119	48	cytoplasm	Hypothetical protein	CAC2557 / <i>Clostridium acetobutylicum</i>	2E-13 (34)	XCC0996	AAK80508
RL059	-	56,768	58,303	511	59	cytoplasm	Transposase	PP4439 / ISPpu14 ORF3 / <i>P. putida</i> KT2440	0.0 (90)		AAN70015
RL060	-	58,367	58,702	111	63	cytoplasm	Transposase	PP4438 / ISPpu14 ORF2 / <i>P. putida</i> KT2440	2E-59 (98)		AAN70014
RL061	-	58,753	59,019	88	57	cytoplasm	Transposase	PP4437 / ISPpu14 ORF1 / <i>P. putida</i> KT2440	7E-39 (88)		AAN70013
RL062	-	59,276	60,502	408	46	cytoplasm	Hypothetical protein	No significant similarity			
RL063	-	60,826	63,075	749	63	inner membrane	Plasmid-related protein	XAC2243 / <i>X. axonopodis</i> pv. citri 306	0.0 (69)		AAM37096
RL064	-	63,180	64,631	483	64	cytoplasm	Plasmid-related protein	XAC2242 / <i>X. axonopodis</i> pv. citri 306	E-133 (65)		AAM37095
RL065	-	64,661	65,266	201	63	cytoplasm	Hypothetical protein	XAC2241 / <i>X. axonopodis</i> pv. citri 306	1E-54 (56)		AAM37094
RL066	-	65,358	65,612	84	60	cytoplasm	Hypothetical protein	XF1757 / <i>X. fastidiosa</i> 9a5c	9E-09 (40)	XAC2240	AAF54566

FIG. 34E

RL067	-	65,680	66,042	120	58	cytoplasm		Hypothetical protein	C77 / <i>P. aeruginosa</i> C	8E-23 (42)	XAC2239	AAN62171
RL068	-	66,112	66,387	91	63	cytoplasm		Hypothetical protein	No significant similarity			
RL069	-	66,384	67,073	229	60	inner membrane		Hypothetical protein	XF1760 / <i>X. fastidiosa</i> 9a5c	3E-50 (53)	XAC2237	AAF84569
RL070	-	67,077	67,427	116	62	cytoplasm		Hypothetical protein	No significant similarity			
RL071	-	67,663	68,370	235	53	cytoplasm		Hypothetical protein	SG91 / <i>P. aeruginosa</i> SG17M	6E-35 (40)		AAN62312
RL072	-	68,855	69,091	78	52	cytoplasm		Hypothetical protein	STY4535 / <i>S. enterica</i> subsp. <i>enterica</i> ser. Typhi CT18	7E-11 (54)	XF1772, XAC2217	CAD09314
RL073	+	69,111	69,377	88	57	cytoplasm		Hypothetical protein	No significant similarity			
RL074	-	69,426	69,965	179	57	cytoplasm		Hypothetical protein	XF1761 / <i>X. fastidiosa</i> 9a5c	1E-28 (52)	XAC2236, STY4534, XCC2093	AAF84570
RL075	-	70,626	71,192	188	58	cytoplasm		Hypothetical protein	No significant similarity			
RL076	+	71,191	71,835	214	60	cytoplasm		Hypothetical protein	No significant similarity			
RL077	-	72,107	72,544	145	68	outer membrane and periplasm	<i>pilM2</i>	Type IV B pilus / adhesion and and protein secretion	PilM / <i>S. typhimurium</i> (plasmid R64)	5E-04 (21)	PA4199, STY4540	BAB91693
RL078	-	72,573	73,901	442	63	outer membrane and periplasm	<i>pilV2</i>	Type IV B pilus / adhesion and and protein secretion	shufflon A' / <i>E. coli</i> (plasmid R64)	3E-65 (38)	STY4550, XAC2151	C26421
RL079	-	73,906	74,847	313	65	inner membrane	<i>pilT2</i>	Type IV B pilus / putative peptidase / adhesion and and protein secretion	BfpF / <i>E. coli</i> (plasmid pB171)	8E-25 (27)	XAC2923, XCC2754, PA0396, PA0395	BAA84845
RL080	-	74,844	75,374	176	60	outer membrane and periplasm	<i>pilS2</i>	Type IV B pilus / adhesion and and protein secretion	PIIS / <i>Shigella sonnei</i> P9 (plasmid Collb-P9)	5E-15 (29)	STY4547	BAA75180
RL081	-	75,396	76,475	359	55	inner membrane	<i>pilR2</i>	Type IV B pilus / adhesion and and protein secretion	PilR / <i>S. typhimurium</i> (plasmid R64)	4E-51 (33)	BfpE, STY4546, XAC0697, XCC3423 (Type II), PA2676, PA3102, STY0164	BAB91688

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September 12, 2003 Serial No.: Not Yet Assigned

Page 79 of 118

Customer No.: 21559

FIG. 34F

RL082	-	76,475	78,055	526	63	inner membrane	<i>pilQ2</i>	ATPase / Type IV B pilus / adhesion and protein secretion	PilQ / <i>S. typhimurium</i> (plasmid R64)		STY4545, bfpD, gspE (Type II)	BAB91689
RL083	-	78,064	78,597	177	69	outer membrane and periplasm	<i>pilP2</i>	Type IV B pilus / adhesion and protein secretion	PilP / <i>S. typhimurium</i> (plasmid R64)	1E-06 (34)	STY4544	BAB91690
RL084	-	78,587	79,912	441	64	inner membrane	<i>pilO2</i>	Type IV B pilus / adhesion and protein secretion	PilO / <i>S. enterica</i> subsp. <i>enterica</i> ser. Dublin	3E-15 (21)	STY4543	AAF81211
RL085	-	79,916	81,538	540	63	outer membrane	<i>pilN2</i>	Secretin / Type IV B pilus / adhesion and protein secretion	PilN / <i>E. coli</i> K-12 (plasmid R721)	6E-84 (35)	bfpB, PA1382 (Type II), XF1527, pefD (general secretion pathway protein)	BAB12647
RL086	-	81,627	82,751	374	66	outer membrane and periplasm	<i>pilL2</i>	Type IV B pilus / adhesion and protein secretion	STY4539 / <i>S. enterica</i> subsp. <i>enterica</i> ser. Typhi CT18	1E-48 (34)		CAD09317
RL087	+	83,023	83,349	108	40	inner membrane		Colicin immunity protein	PA0984 / <i>P. aeruginosa</i> PAO1	4E-55 (94)		AAG04373
RL088	-	83,374	84,870	498	47	inner membrane		Colicin-like toxin (pyocin S5)	PA0985 / <i>P. aeruginosa</i> PAO1	0.0 (97)	PA1150	AAG04374
RL089	-	85,558	87,531	657	64	inner membrane		DNA Helicase	DNA helicase / <i>Dictyostelium discoideum</i> AX14	4E-28 (24)	PA0799	AAO51158
RL090	-	87,528	89,417	629	60	cytoplasm		Hypothetical protein	No significant similarity			
RL091	-	89,551	90,021	156	44	cytoplasm		Similar to luminal binding protein	Rlorf142 / <i>Rhizobium rhizogenes</i> MAFF03-01724	5E-29 (41)		BAB16261
RL092	-	90,095	92,020	641	62	cytoplasm	<i>topA</i>	Topoisomerase I	XF-a0003 / <i>X. fastidiosa</i> 9a5c (plasmid pXF51)	E-151 (43)	PA3011, XF1776, XF0920, STM1298, XCC3755, STY1336	AAF85572
RL093	+	92,340	92,573	77	58	cytoplasm		Hypothetical protein	No significant similarity			
RL094	-	92,712	92,957	81	59	cytoplasm		Hypothetical protein	No significant similarity			
RL095	-	93,421	93,891	156	61	cytoplasm	<i>ssb</i>	Single-stranded DNA binding protein	C102 / <i>P. aeruginosa</i> C	6E-32 (53)	XAC2211, PA4232	AAN62318
RL096	-	93,905	94,438	177	60	cytoplasm		Hypothetical protein	XF1779 / <i>X. fastidiosa</i> 9a5c	8E-39 (53)	XAC2210	AAF84587
RL097	-	94,444	95,172	242	63	cytoplasm		Hypothetical protein	ORF6 / <i>Pseudomonas</i> sp. B13	7E-66 (52)	STY4529	CAD60670

FIG. 34G

RL098	-	95,442	95,681	79	46	inner membrane		Hypothetical protein	No significant similarity				
RL099	-	95,672	95,947	91	60	cytoplasm		Hypothetical protein	No significant similarity				
RL100	-	95,944	97,269	441	60	cytoplasm		Hypothetical protein	ORF5 / <i>Pseudomonas</i> sp. B13	1E-68 (39)	XAC2208, STY4528	CAD60669	
RL101	-	97,266	98,033	255	58	cytoplasm		Hypothetical protein	ORF4 / <i>Pseudomonas</i> sp. B13	2E-33 (40)	XF1782, STY4526	CAD60668	
RL102	-	98,061	99,800	579	58	cytoplasm		Hypothetical protein	SG102 / <i>P. aeruginosa</i> SG17M	E-116 (43)	XF1783, XF1784, STY4523	AAN62323	
RL103	-	99,797	100,051	84	60	cytoplasm		Hypothetical protein	No significant similarity				
RL104	-	100,048	101,064	338	64	cytoplasm		Hypothetical protein	PA3849 / <i>P. aeruginosa</i> PAO1	0.0 (96)	STY2463, STY1607, STM2226	AAG07236	
RL105	-	101,064	101,297	77	65	cytoplasm		Hypothetical protein	No significant similarity				
RL106	-	101,290	101,784	164	62	cytoplasm		Hypothetical protein	No significant similarity				
RL107	-	101,777	102,034	85	60	cytoplasm		Hypothetical protein	No significant similarity				
RL108	-	102,031	102,558	175	62	cytoplasm		Hypothetical protein	No significant similarity				
RL109	-	102,745	104,107	pseudogene	61		<i>dnaB</i>	DNA replication and recombination	Replicative DNA helicase DnaB / <i>Pseudomonas</i> sp. SLT2001 (plasmid pQBR55)	E-130 (60)	PA4931, STY4442, XAC1477, XF0361, XCC1434	CAD13464	
RL110	-	104,282	104,989	235	60	cytoplasm		Putative phage protein	FNV0875 / <i>Fusobacterium nucleatum</i> subsp. <i>vincentii</i> 49256	2E-21 (34)		EAA24090	
RL111	-	104,986	105,687	233	64	cytoplasm		Hypothetical protein	STY1595 / <i>S. enterica</i> subsp. <i>enterica</i> ser. Typhi CT18	1E-16 (32)		CAD01841	
RL112	-	105,687	106,373	228	64	inner membrane		Hypothetical protein	No significant similarity				
RL113	-	106,516	107,013	165	60	outer membrane and periplasm		Hypothetical protein	PA2226 / <i>P. aeruginosa</i> PAO1	2E-32 (47)		AAG05614)	
RL114	-	107,010	107,690	226	58	cytoplasm		Hypothetical protein	ORF50 / <i>P. putida</i> plasmid pDTG1	2E-25 (35)		AF491307	
RL115	-	107,687	108,616	309	60	inner membrane	<i>soj</i>	Chromosome partitioning	Soj / <i>P. aeruginosa</i> SG17M	1E-84 (90)	XF1785, XAC2205, STY4521	AAG02083	
		108,700	108,757		47				<i>attR</i> / <i>P. aeruginosa</i> PAO1				

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September 12, 2003 Serial No.: Not Yet Assigned

Page 81 of 118

Customer No.: 21559

FIG. 35

Alignment: No_2 - embi[AL039136]HSM003612

Homo sapiens mRNA; EST DKFZp566K094_r1(from clone DKFZp566)

Q:18 DQTCDNLSQNPPHLLLRLLDHWGDPAGCWSLGQTYSGHLYLPYCRELHKCSLCAHRNWH SEQ ID NO:230
DQTCDNLSQNPPHLLLRLLDHWGDPAGCWSLGQTYSGHLYLPYCRELHKCSLCAHRNWH SEQ ID NO:231
H:29 DQTCDNLSQNPPHLLLRLLDHWGDPAGCWSLGQTYSGHLYLPYCRELHKCSLCAHRNWH SEQ ID NO:232

HYCCLWPVWMLCYMSW 93
HYCCLWPVWMLCYMSW
HYCCLWPVWMLCYMSW 256

FIG. 36

Alignment: No_8 - embi[M79137]HSXT01285

EST01285 Subtracted Hippocampus, Stratagene (cat.#936205)H

Q:18 QVQHPPLCLLDQHQQECIPPCLPPDHLQDPQHFPFLPDHHVPHLVVLIQPQLCRALAP SEQ ID NO:233
75
QVQHP .CLLDQHQQECIPPCLPPDHLQDPQHFPFLPDHHVPHLVVLIQPQLCRALAP SEQ ID NO:234
H:43 QVQHPXXCLLDQHQQECIPPCLPPDHLQDPQHFPFLPDHHVPHLVVLIQPQLCRALAP SEQ ID NO:235
216

FIG. 37

Alignment: No47-swissnew[p35555]FBN1_HUMAN

FIBRILLIN 1 PRECURSOR.//:swiss[P35555]FBN1_HUMAN FIBRILLIN1

PRECURSOR.//:trernbl [L13923]HSFIBRLLN_1 product: "fibrillin"; Homo sapiens fibrillin

mRNA, complete cds.//:gp[L13923]306746 product: "fibrillin"; Homo sapiens fibrillin mRNA,complete cds.

Q: 18 CGGASCHNTLGSYKCMCPAGFQYEQFSGGCQDINECGSAQAPCSYGCSNTEGGYLCGCPP SEQ ID NO:236
CGGASCHNTLGSYKCMCPAGFQYEQFSGGCQDINECGSAQAPCSYGCSNTEGGYLCGCPP SEQ ID NO:237
H:2617 CGGASCHNTLGSYKCMCPAGFQYEQFSGGCQDINECGSAQAPCSYGCSNTEGGYLCGCPP SEQ ID NO:238

GYFRIGQGHCVSGMGMRGNPEPPVSGEMDDNSLSPEACYECKINGYPKRGRKRRSTNET SEQ ID NO:236
GYFRIGQGHCVSGMGMRGNPEPPVSGEMDDNSLSPEACYECKINGYPKRGRKRRSTNET SEQ ID NO:237
GYFRIGQGHCVSGMGMRGNPEPPVSGEMDDNSLSPEACYECKINGYPKRGRKRRSTNET SEQ ID NO:238

DASNIEDQSETEANVSLASWDVEKTAIFAFNISHV-NKVRIL 178 SEQ ID NO:236
DASNIEDQSETEANVSLASWDVEKTAIFAFNISHV NKVRIL SEQ ID NO:237
DASNIEDQSETEANVSLASWDVEKTAIFAFNISHVSNKVRIL 2778 SEQ ID NO:238

FIG. 38

Alignment: No56-trembl[AF088916]AF088916_1

gene:"EMI";product:"elastin microfibril interfase located protein"; Homo sapiens elastin microfibril interfase located protein (EMI) gene, complete cds.

//:trembl [AF088916]AF088916_1 product: "emilin precursor"; Homo sapiens emilin precursor, mRNA, complete cds and 3' UTR.//:gp[AF088916]5353510 product:"emilin precursor"; Homo sapiens emilin precursor, mRNA, complete cds and 3'UTR.

//:gpnew[AF162780]6693840 gene:"EMI"; product: "elastin microfibril interfase located protein; Homo sapiens elastin microfibril interfase located protein (EMI) gene, complete cds.

Q: 7 DGDVYNPSTGVFTAPYDGRYLITATLTPERDAYVEAVLSVSNASVAQLHTAGYRREFLEY SEQ ID NO:239
 DG..Y:P.TGVFTAP. GRYL::A.LT .R.. VEAVLS SN..VA::...GY. E LE SEQ ID NO:240
 H:896 DGGYYDPETGVFTAPLAGRYLLSAVLTGHRHEKVEAVLSRSNQGVARVDSGGYEPEGLE- SEQ ID NO:241

HRPPGALHTCGGP-GAFHLIVHLKAGDAV 94 SEQ ID NO:239
 ::P .. :.. G. G.F.LI: L:AGD.V SEQ ID NO:240
 NKPVAESQSPGTLGVFSLILPLQAGDTV 983 SEQ ID NO:241

gene:"EMI";product:"elastin microfibril interfase located protein"; Homo Sapiens elastin microfibril interfase located protein (EMI) gene, complete cds.

//:trembl [AF088916]AF088916_1 product: "emilin precursor"; Homo sapiens emilin precursor, mRNA, complete cds and 3' UTR.//:gp[AF088916]5353510 product:"emilin precursor" Homo sapiens emilin precursor, mRNA, complete cds and 3'UTR.

//:gpnew[AF162780]6693840 gene:"EMI"; product: "elastin microfibril interfase located protein", Homo sapiens elastin microfibril interfase located protein (EMI) gene, complete cds

Q: 7 DGDVYNPSTGVFTAPYDGRYLITATLTPERDAYVEAVLSVSNASVAQLHTAGYRREFLEY SEQ ID NO:242
 DG..Y:P.TGVFTAP. GRYL::A.LT .R.. VEAVLS SN..VA::...GY. E LE SEQ ID NO:243
 H:896 DGGYYDPETGVFTAPLAGRYLLSAVLTGHRHEKVEAVLSRSNQGVARVDSGGYEPEGLEN SEQ ID NO:244

HRPPGALHTCGGPGAFHLIVHLKAGDAV 94 SEQ ID NO:242
 . . : : G. G.F.LI: L:AGD.V SEQ ID NO:243
 KPVAESQSPGTLGVFSLILPLQAGDTV 983 SEQ ID NO:244

FIG. 39

Alignment: No59-pironly[A35763]A35763

unnamed ORF;P.lividus 2-alpha collagen(COLL2-alpha) mRNA, complete cds.
 //:pironly[A35763]A35763 collagen alpha 2 chain-sea urchin(Paracentrotus lividus)
 (fragment)//:gp[J05422]159962 unnamed ORF;P.lividus 2-alpha collagen(COLL2-alpha)
 mRNA, complete cds.

Q: 92 GENGSSGSQAPLQGLRGIFGLWGRRSRARFCGPR-PVARLGGGTSAGRELGL 142 SEQ ID NO:245
 GE G.SG...P QG:RGI G: G.... GPR P . GGG S G.. GL SEQ ID NO:246
 H: 718 GEPGPSGENGP-QGVRGIPGVVGENGKTGRGGPRGPPGLRGGGSRGERGGL 768 SEQ ID NO:247

unnamed ORF;P.lividus 2-alpha collagen(COLL2-alpha) mRNA, complete cds. //:pironly[A35763]A35763
 collagen alpha 2 chain-sea urchin(Paracentrotus lividus) (fragment)//:gp[J05422]159962 unnamed
 ORF;P.lividus 2-alpha collagen(COLL2-alpha) mRNA, complete cds.

Q: 92 GENGSSGSQAPLQGLRGIFGLWGRRSRARFCGPR-PVARLGGGTSAGRELGL 142 SEQ ID NO:248
 GE G.SG...P QG:RGI G: G.... GPR P . GGG S.G.. GL SEQ ID NO:249
 H: 718 GEPGPSGENGP-QGVRGIPGVVGENGKTGRGGPRGPPGLRGGGSRGERGGL 768 SEQ ID NO:250

FIG. 40

Alignment: No60/63-swiss[P20062]TCO2_HUMAN

TRANSCOBALAMIN II PRECURSOR.//:treml[M60396]HSTCII_1gene:"TCN2"; product:
"transcobalamin II"; Human transcobalamin II (TCII) mRNA, complete cds.
//:gp[M60396]339196 gene: "TCN2"; product: "transcobalamin II"; Human transcobalamin II (TCII)
mRNA, complete cds.

Q: 8	VEPFHQGHHSVDTAAMAGLAFTCLKRSNFPNPGRRQRITMAIRTVREEILKAQTPEGHFGN	SEQ ID NO:251
	VEPFHQGHHSVDTAAMAGLAFTCLKRSNFPNPGRRQRITMAIRTVREEILKAQTPEGHFGN	SEQ ID NO:252
H:183	VEPFHQGHHSVDTAAMAGLAFTCLKRSNFPNPGRRQRITMAIRTVREEILKAQTPEGHFGN	SEQ ID NO:253
	VYSTPLALQFLMTSPMPGAELGTACLKARVALLASLQDGAFQNALMISQLLPVLNHKTYI	SEQ ID NO:251
	VYSTPLALQFLMTSPMPGAELGTACLKARVALLASLQDGAFQNALMISQLLPVLNHKTYI	SEQ ID NO:252
	VYSTPLALQFLMTSPMPGAELGTACLKARVALLASLQDGAFQNALMISQLLPVLNHKTYI	SEQ ID NO:253
	DLIFPDCLAPRVMLEPAA	145 SEQ ID NO:251
	DLIFPDCLAPRVMLEPAA	SEQ ID NO:252
	DLIFPDCLAPRVMLEPAA	320 SEQ ID NO:253

TRANSCOBALAMIN II PRECURSOR.//:treml[M60396]HSTCII_1gene:"TCN2"; product:
"transcobalamin II"; Human transcobalamin II (TCII) mRNA, complete cds.
//:gp[M60396]339196 gene: "TCN2"; product: "transcobalamin II"; Human transcobalamin II
(TCII) mRNA, complete cds.

Q: 8	VEPFHQGHHSVDTAAMAGLAFTCLKRSNFPNPGRRQRITMA	47 SEQ ID NO:254
	VEPFHQGHHSVDTAAMAGLAFTCLKRSNFPNPGRRQRITMA	SEQ ID NO:255
H:183	VEPFHQGHHSVDTAAMAGLAFTCLKRSNFPNPGRRQRITMA	222 SEQ ID NO:256

TRANSCOBALAMIN II PRECURSOR.//:treml[M60396]HSTCII_1gene:"TCN2";product:
"transcobalamin II"; Human transcobalamin II (TCII) mRNA, complete cds.
//:gp[M60396]339196 gene: "TCN2"; product: "transcobalamin II"; Human transcobalamin II
(TCII) mRNA, complete cds.

Q: 8	VEPFHQGHHSVDTAAMAGLAFTCLKRSNFPNPGRRQRITMA	47 SEQ ID NO:257
	VEPFHQGHHSVDTAAMAGLAFTCLKRSNFPNPGRRQRITMA	SEQ ID NO:258
H:183	VEPFHQGHHSVDTAAMAGLAFTCLKRSNFPNPGRRQRITMA	222 SEQ ID NO:259

FIG. 41

Alignment: No65-swissnew[P23142]FBL1_HUMAN

FIBULIN-1 PRECURSOR.//:swiss[P37888]FBLD_HUMAN FIBULIN-1, ISOFORM D
 PRECURSOR.//:treml[U01244]HS2444_1 product: "fibulin-1D"; Homo sapiens fibulin-1D
 mRNA, complete cds.//:gp[U01244]1621019 product: "fibulin-1D"; Homo sapiens fibulin-1D
 mRNA, complete cds.

```
Q: 18 RNCQDIDECVTGIHNCSINETCFNIQGGFRCLAFECPENYRRSAATLQQEKDQTVRCIKS SEQ ID NO:260
      RNCQDIDECVTGIHNCSINETCFNIQG.FRCLAFECPENYRRSAATLQQEKDQTVRCIKS SEQ ID NO:261
H:521 RNCQDIDECVTGIHNCSINETCFNIQGAFRCLAFECPENYRRSAATLQQEKDQTVRCIKS SEQ ID NO:262

      CRPNDVTCVFDPVHTISHTVISLPTFREFTRPEEIIFLRAITPPHPASQANIIDITEGN SEQ ID NO:260
      CRPNDVTCVFDPVHTISHTVISLPTFREFTRPEEIIFLRAITPPHPASQANIIDITEGN SEQ ID NO:261
      CRPNDVTCVFDPVHTISHTVISLPTFREFTRPEEIIFLRAITPPHPASQANIIDITEGN SEQ ID NO:262

      LRDSFDIIKRYMDGMTVGIRR 158 SEQ ID NO:260
      LRDSFDIIKRYMDGMTVG: R SEQ ID NO:261
      LRDSFDIIKRYMDGMTVGVR 661 SEQ ID NO:262
```


FIG. 42

Alignment: 80 - trembl[AF045447]AF045447_1

gene: "DPC4"; product: "deleted in pancreatic carcinoma"; Homo sapiens deleted in pancreatic carcinoma (DPC4) gene, exon 11 partial sequence and complete cds. //:trebl[U44378]HS443781_1 gene: "DPC4"; product: "Dpc4"; Human homozygous deletion target in pancreatic carcinoma (DPC4) mRNA, complete cds. //:pironly[S71811]S71811 probable transcription regulator MAD-4-human//:gp[AF045447]2865657 gene: "DPC4"; product: "deleted in pancreatic carcinoma"; Homo sapiens deleted in pancreatic carcinoma (DPC4) gene, exon 11 partial sequence and complete cds. //:gp[U44378]1163234 gene: "DPC4"; product: "Dpc4"; Human homozygous deletion target in pancreatic carcinoma (DPC4) mRNA, complete cds.

```

Q:  6  PGSRIRGRVDTLQXNAPXXMMVKDEYVHDFEGQPXLXTEGHXIQTIQHPPXNRAXTETYX  SEQ  ID  NO:263
      PG  :.G    TLQ.NAP..MMVKDEYVHDFEGQP.L.TEGH.IQTIQHPP.NRA.TETY.  SEQ  ID  NO:264
H:139  PGIDL SGL--TLQSNAPSSMMVKDEYVHDFEGQPSLSTEGHSIQTIQHPPSNRASTETYS  SEQ  ID  NO:265

      TPALLAPXEXNATXTANFPNIPVAXTXQPAXILGGXHXEGLLQIAXGPQPGQQQNGFTGQ  SEQ  ID  NO:263
      TPALLAP.E.NAT.TANFPNIPVA.T.QPA.ILGG.H.EGLLQIA.GPQPGQQQNGFTGQ  SEQ  ID  NO:264
      TPALLAPSESNATSTANFPNIPVASTSQPASILGGSHSEGLLQIASGPQPGQQQNGFTGQ  SEQ  ID  NO:265

      PATYHHNXTTTWTGXRTAPYTPNLPHHQKG 155  SEQ  ID  NO:263
      PATYHHN.TTTWTG.RTAPYTPNLPHHQ.G    SEQ  ID  NO:264
      PATYHHNSTTTWTGSRTAPYTPNLPHHQNG 286  SEQ  ID  NO:265

```

FIG. 43

Alignment: No86 - trembl[D32210]D32210 _1

gene: "Notch2"; product:"cell surface protein"; Mus musculus (Notch2) mRNA, complete cds.
 //:gp[D32210]2373395 gene: "Notch2"; product:"cell surface protein"; Mus musculus
 (Notch2) mRNA, complete cds.

Q:81 MPALRPALLWALLALWLCCATPAHALQCRDGYEPCVNEGMCVITYHNGTGYCKCP-GFLGE SEQ ID NO:266
 MP LRPA.L ALL LWLC A PAHALQCR.G.EPCVNEG.CVITYHNGTG:C:CP GFLGE SEQ ID NO:267
 H: 1 MPDLRPAALRALLWLWLCGAGPAHALQCRGGQEPVNEGTCVITYHNGTGFCRCPEGFLGE SEQ ID NO:268

YCQHR-PCEKNRCGDPSTC 157 SEQ ID NO:266
 YCQHR PCEKNRC : .TC SEQ ID NO:267
 YCQHRDPCEKNRCQNGGTC 79 SEQ ID NO:268

FIG. 44

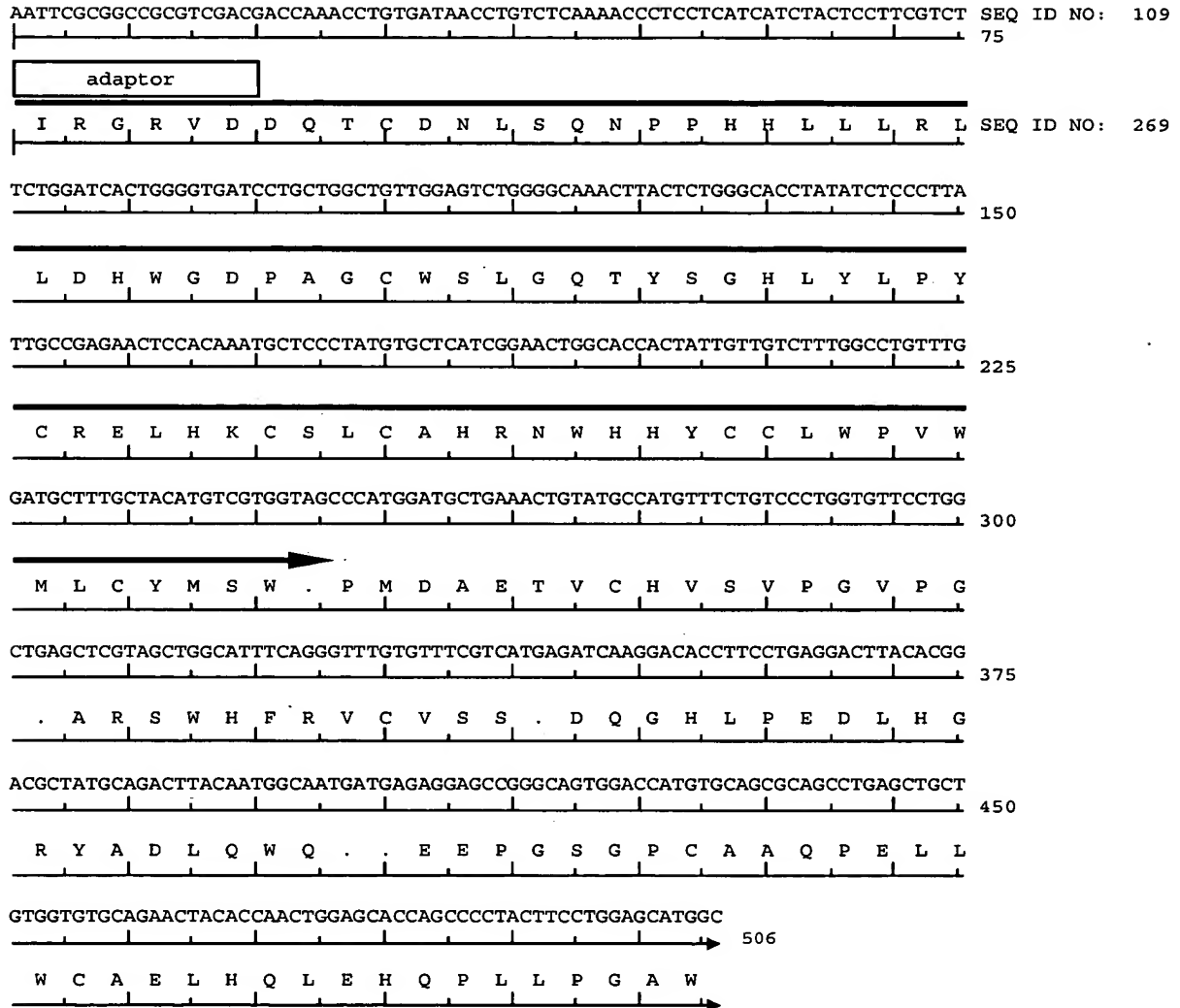


FIG. 45

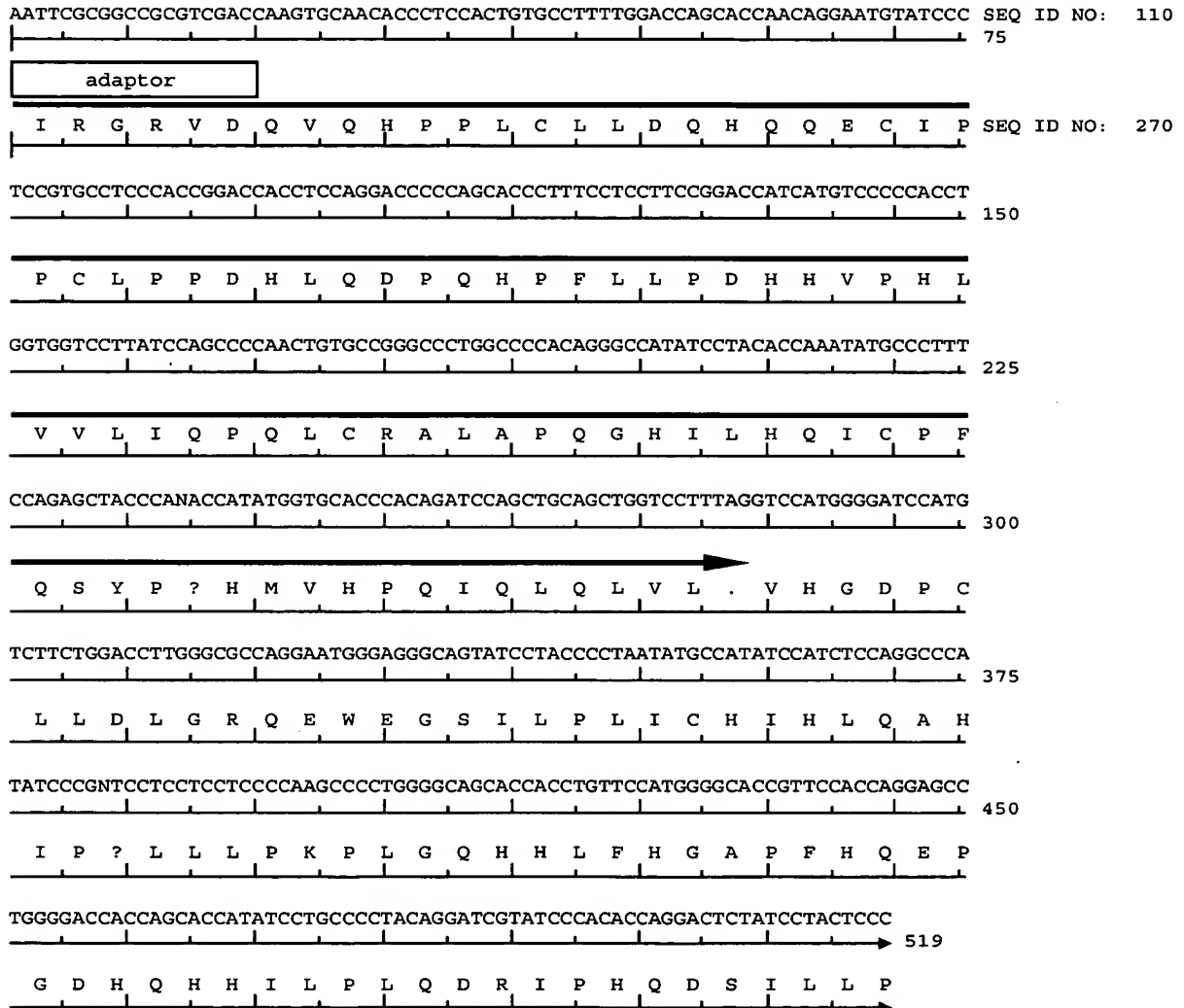


FIG. 46

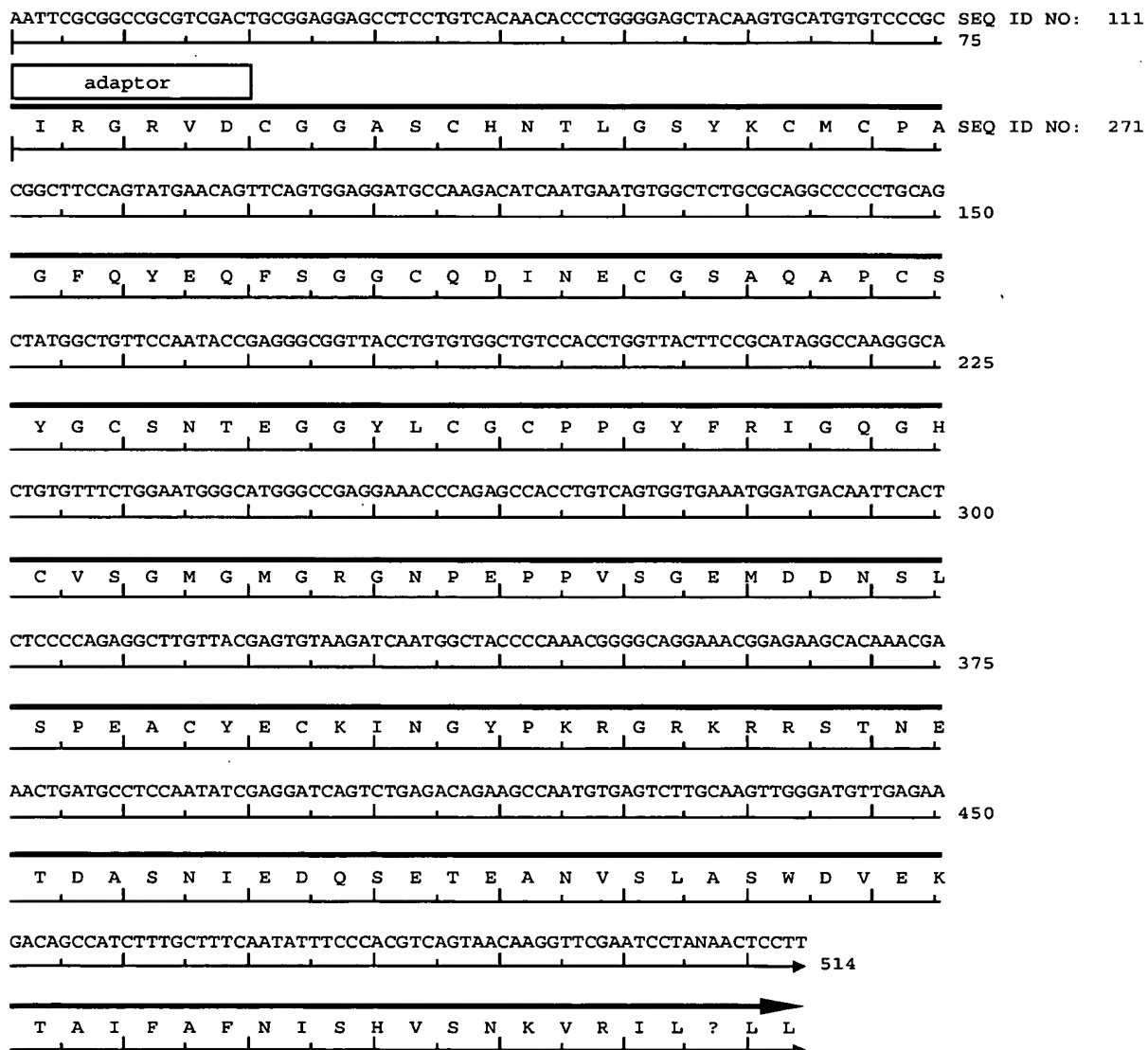


FIG. 47

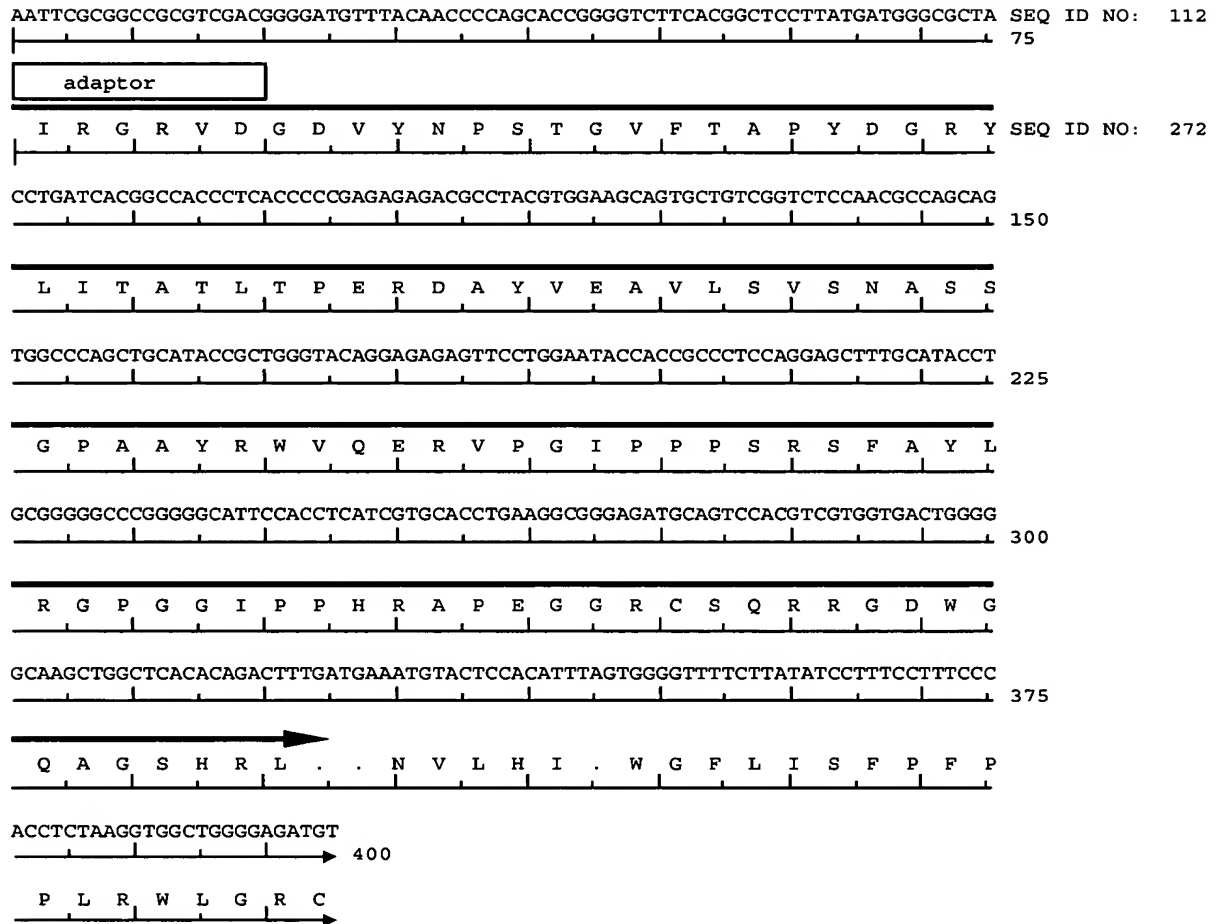


FIG. 48

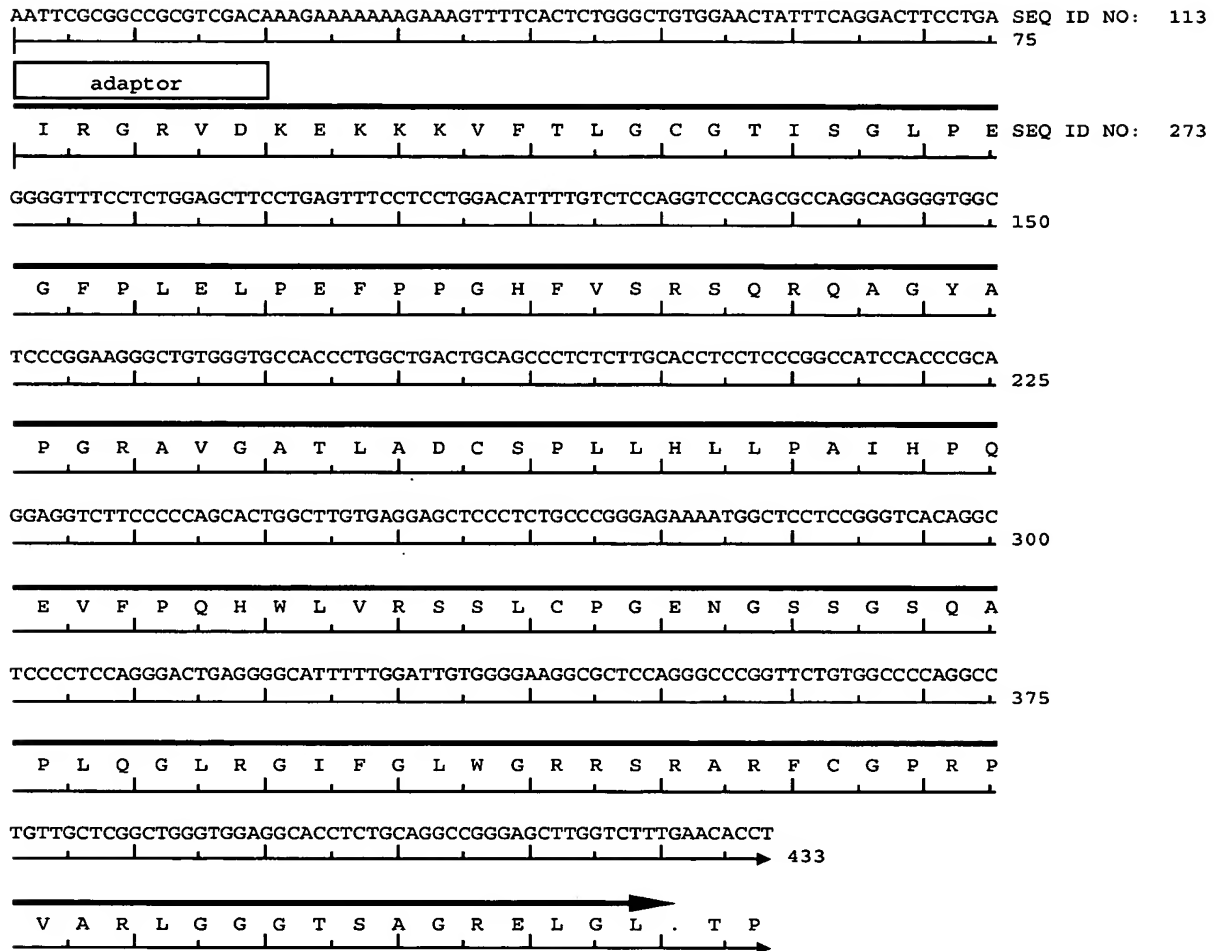


FIG. 49

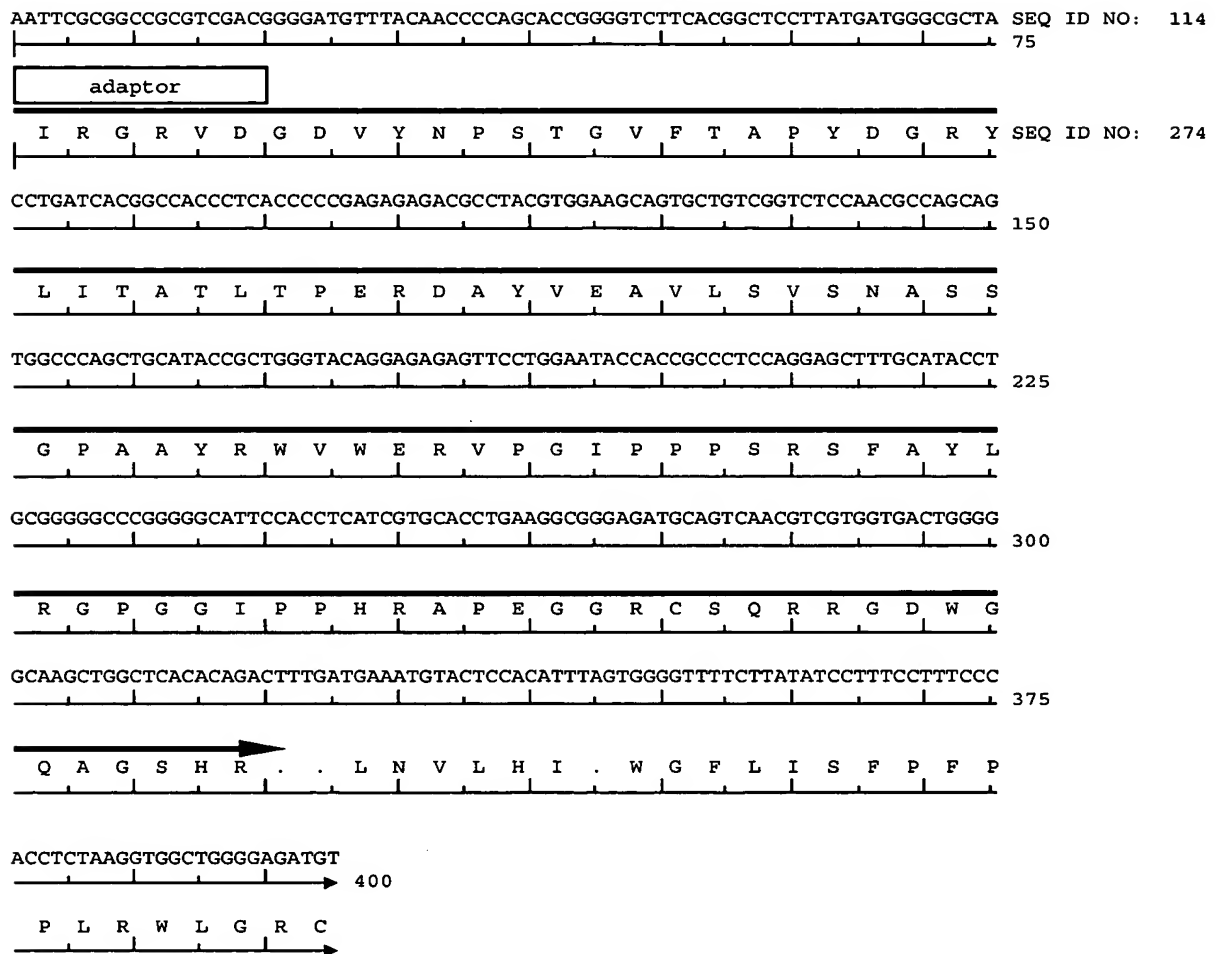


FIG. 50

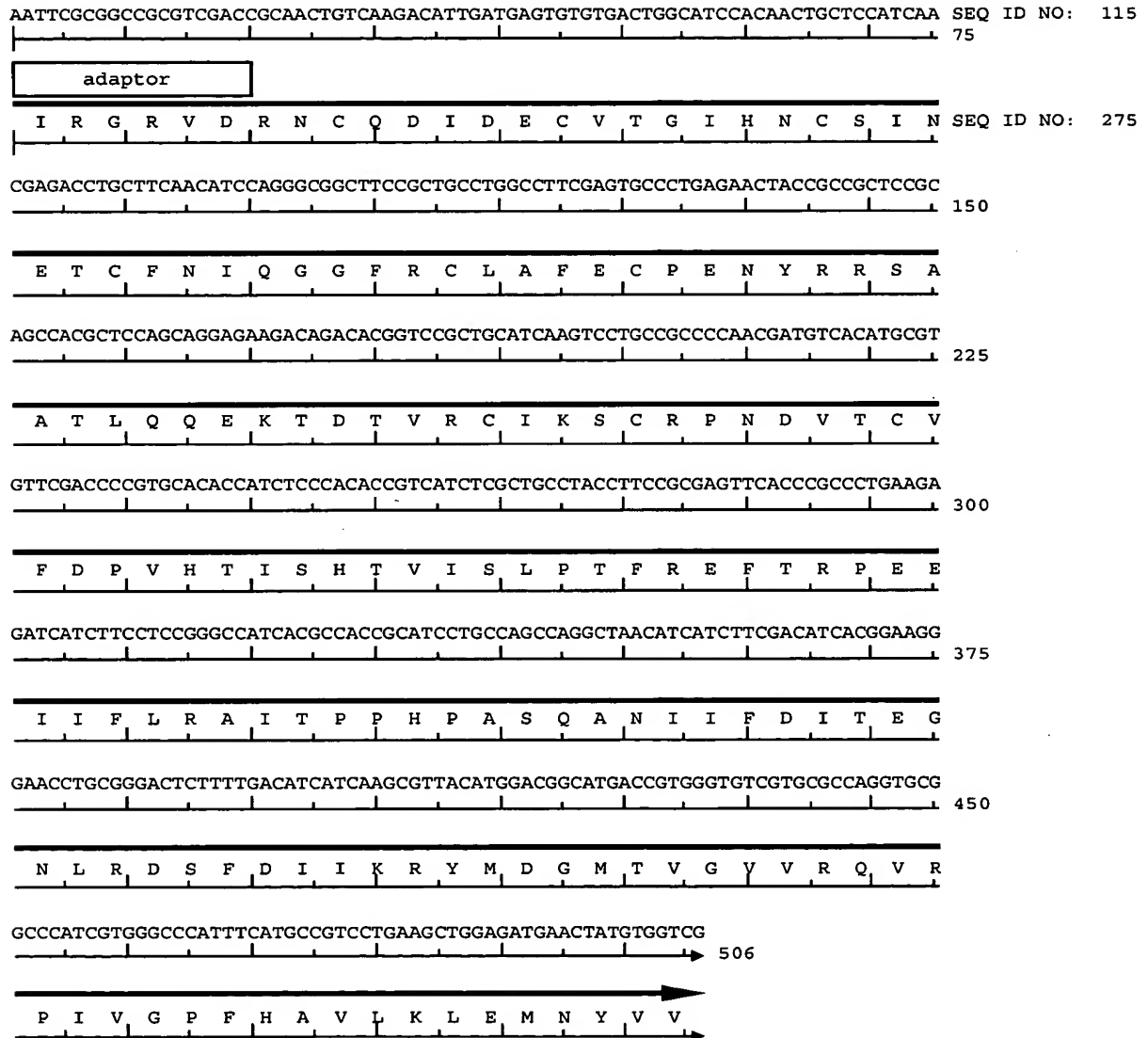


FIG. 51

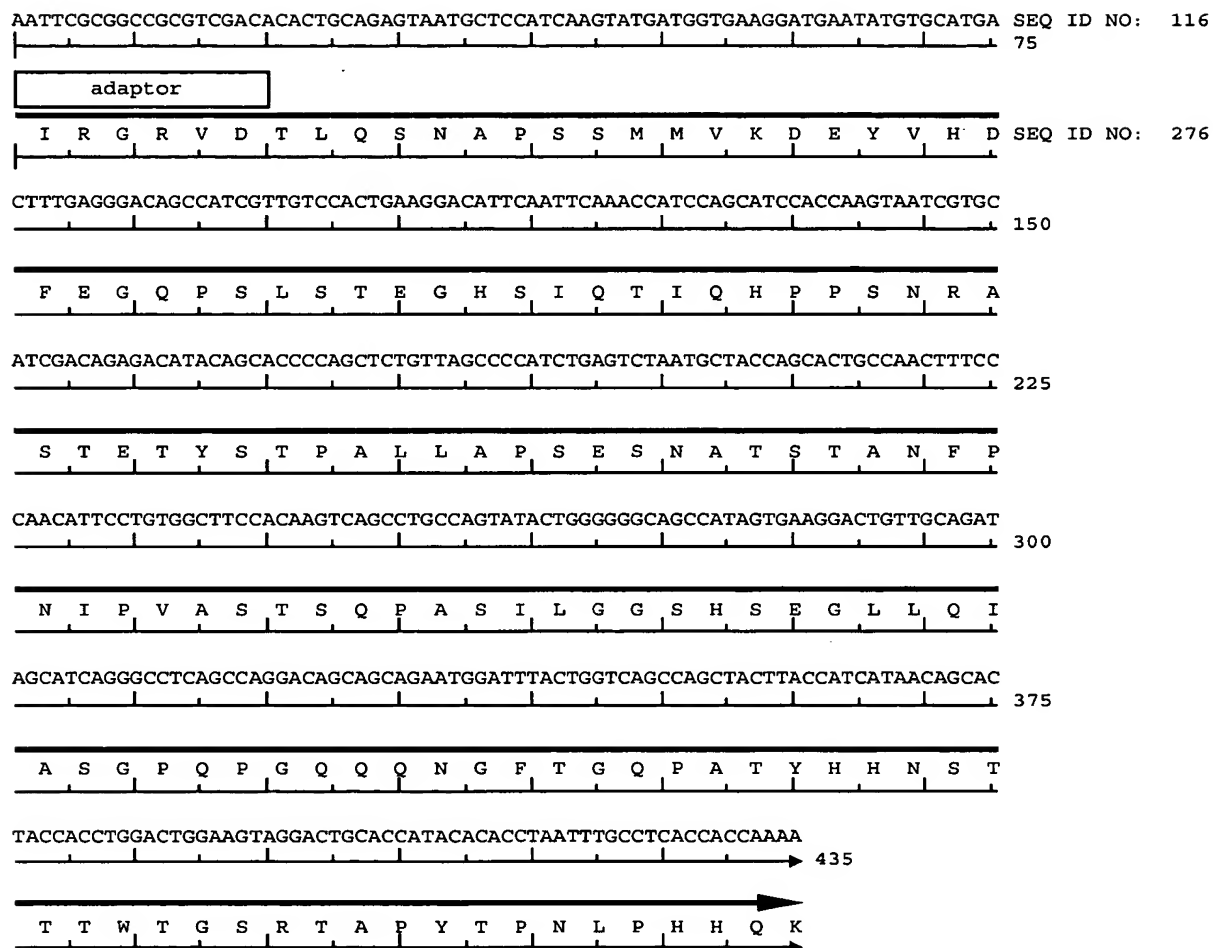


FIG. 52

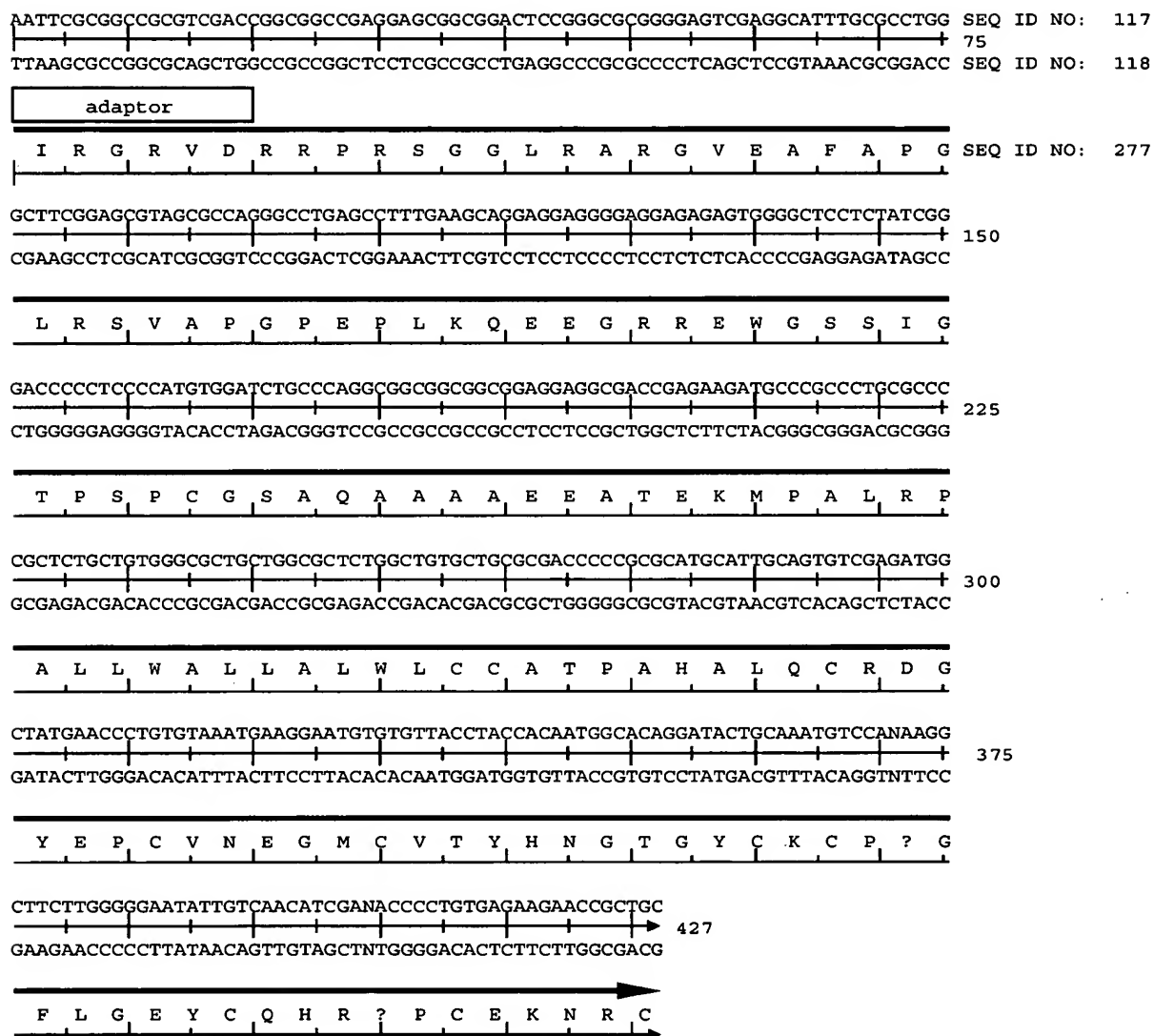


FIG. 53

Pathogenicity Island probe # (bp#)	<i>P. aeruginosa</i> strains that gave Positive signal with the specific probe used	<i>P. aeruginosa</i> strains that gave Negative signal with the specific probe used
3 (25562-26456)	PA14, CF2, CF6, CF26, CF29	PAO1, PAK, CF1, CF3, CF4, CF5, CF27, CF28, CF30, CF32
4 (61181-63607)	PA14, CF2, CF6, CF26, CF29	PAO1, PAK, CF1, CF3, CF4, CF5, CF27, CF28, CF30, CF32
5 (74933-76117)	PA14, PAO37, CF2, CF6, CF26	PAO1, PAK, CF1, CF3, CF4, CF5, CF27, CF28, CF29, CF30, CF32
6 (84922-86622)	PAO1, PA14, PAO37, CF2, CF6, CF26	PAK, CF1, CF3, CF4, CF5, CF27, CF28, CF29, CF30, CF32
7 (103070-104556)	PA14, PAO37, CF2, CF6, CF26	PAO1, PAK, CF1, CF3, CF4, CF5, CF27, CF28, CF29, CF30, CF32
8 (104799-105545)	PA14, CF2, CF6, CF26	PAO1, PAK, CF1, CF3, CF4, CF5, CF27, CF28, CF29, CF30, CF32

These experiments indicate that at least part of the big island region contained in each probe is present in the *P. aeruginosa* strains that gave positive hybridization signal.

FIG. 54

ORF7 Protein SEQ ID NO: 278

MINSHLLYRLSYRGTSFFQPWTLPVLLDSRLRGAPFYGCARACQPSDPKSFSSFSTSDKTALPL
HAAALSRLPDAHEKAPPKRGFPKPPPKRSGEDDLVAFHLRRDTGTRREFAGQDQLRQRVLDPAL
DGPLQRACAI DRVEADGNQLVQRLLAQFQAQLALGQALAQATELDLG DAGDLLASQRLEHHHFV
DPVDEFRTFVRIDRVHHCGLRLAVAGQLDLRRTEVGGHHHHGVAEVHRTPTVTVGQASVLEHL
EENVEYIRMGLLHLVQQHHRVGLAADRLGQVAAFLEADVARRRADQAGHRVFLHELGHYYPHQR
LLGIEEELGQRLAQLGLAHPGRAEEEEERAARPVRIGEAGARTAHGVGHGDYRLVLADHSPMQLL
LHAQQLLALALEHLRHRDTGPLGNHFGDFLVGHLVAQQVLGLAVLVDHLQAAFQVRDGLVLDA
RHALEVALAPRRHLHLLGLLDLLDLRRALHLGLLGLPDLLLEVGVFALELDDILLQLGQALPGG
FVVFLQLRLALDLQLDQATVETIQFLRLGVDLHADAAGGLVDQVDGLVRQLPIGDVAVRQLGRG
DDRAVGDAHPVVHFI AFLEATEDGDGVFLARFVHQHLLAALQRGILLDVLAILVEGSSTDAVQ
LAARQSRLEHVAGVHGTFRLAGADHGVQFVDEQDDPAFLLAQFVEDRLQAFLELAAELGTGDQR
PHVQQQALVLEAVRHFAVDDALGQALDDGGLADAGFADQHRVVLGPPLQDLGDPADLVVATDH
RVELAFLGALGHVDGVLVQRLARLLDVRVVRFAATQVGHGILQRLARHALAEQQLAEPGVLVH
RGQQYQLAGDELVALLLGQAVSLVEQACEILGQVHVAGRALDLRQRVEFFVEAAAQGGDIEADL
HQQGLDRTALLLEQGGKQVHRLDGRMVMANGQGLGVGERQLQLAGQTVYSHGSSFLL.

FIG. 55

ORF7 nucleotide sequence SEQ ID NO: 119

ATGATTAACAGTCATTTGCTCTACCACTGAGCTATCGCGGAACGTCTTTCTTCCAACCCCTGGAC
GCTTCCGGTGTTGCTGGATTTCGCGTCTCAGAGGCGCGCCATTTTACGGATGCGCGCGGGCATGT
CAACCCTCTGATCCAAAAAGTTTTTCTTCTTTTCCACGAGCGACAAAACGGCCCTTCCACTGC
ATGCGGCAGCGCTCTCGCGCTACCGGACGCCCATGAAAAAGCCCCGCCGAAGCGGGGCTTTCC
CTGTCCGCCCCCGAAGAGGTGAGGCGAAGACGATCTCGTCGCCTTCCACCTTCGCCGAGATACG
GGCGTGCGCCATAGACCGGGTCGAAGCCGACGGCAATCAGCTTGTCCAGCGCCTCCTGGCTCAG
TTCCAAGGCTCAGCTCGCGCTCGGCCAGGCGCTTGCGCAGGCGACCGAGCTGGATCTCGGCGAT
GCCGGCGATCTGCTCGCGAGCCAGCGGCTCGAACACCACCACTTTCGTGATCCGGTTGATGAAT
TCCGGACGGAAGTGCGCATTGACCGCGTCCATCACTGCGGCACGTTGCGCCTCGCGGTGCGCCG
CCAGCTCCTGGATCTGCGCCGAACCGAGGTTGGAGGTGATCACCACCACGGTGTTGCGGAAGTC
CACCGTACGCCCCTGACTGTGCGGTGAGGCGTCCGTCTTCGAGCACCTGGAGGAGAATGTTGAAT
ACATCCGGATGGGCCTTCTCCACCTCGTCCAGCAGCACCAACGAGTAGGGCTTGCGGCGGATCG
CCTCGGTGAGGTAGCCGCTTCTCGAAGCCGACGTAGCCCGGAGGCGCGCCGATCAGGCGGGC
CACCGAGTGTTTCTCCATGAACTCGGACATATCTATCCGCACCAGCGCCTCCTCGGTATCGAAG
AGGAACTCGGCCAGCGCCTTGCACTCGGTCTTGCCCAACCCCGGTGCGGCGGAGGAAGAGGA
ACGAGCCGCTCGGCCGTTGCGATCGGCGAGGCGGCGCGCAACGGCGCACGGCGTTGGACAC
GGCGACTACCGCCTCGTCTGCGCGATCACTCGCCGATGCAGCTCCTGCTCCATGCGCAGCAGC
TTCTCGCGCTCGCCCTCGAGCATCTTCGACACCGGGATACCGGTCCACTTGGAACCACTTTCGG
CGATTTCTCGTTCGGTACCTGTTGCGCAGCAACTGGTTCTCGGTCTTGCCGTGCTGGTTCGACC
ATCTGCAGGCTGCGTTCCAGGTCCGGGATGGTCTGGTACTGGATGCGCGCCATGCTCTCGAGGT
CGCCCTTGCGCCGCGCCGCTCCATCTCCTGCTTGGCCTGCTCGATCTTCTGCTGGATCTGCGC
CGAGCCCTGCACCTCGGCCTTCTCGGACTTCCAGATCTCCTCGAGGTGCGCGTATTTCGCGCTCG
AGCTTGACGATATCCTCCTCCAGCTTGGCCAGGCGCTTCTGGTGGCTTCGTGCTCTTCTCTCT
TCAGCGCCTCGCGCTCGATCTTCAGCTGGATCAGGCGACGGTTCGAGACGATCCAGTTCTCCTCGG
CTTGAGTTCGATCTCCATGCGGATGCGGCTGGCGGCCTCGTCGATCAGGTTCGATGGCCTTGTC
GGCAGTTGCCGATCGGTGATGTAGCGGTGCGACAGCTTGGCCGCGGCGATGATCGCGCCGTCGG
TGATGCTCACCCGTGGTGCATTCATAGCGTTCTTGAGGCCACGGAGGATGGCGATGGTGTCT
TCCTCGCTCGGTTTCGTCCACCAGCACCTTCTGGAAGCGGCGCTCCAGCGCGGCATCCTTCTCGA
TGTAAGTGGCGATACTCGTCGAGGGTAGTAGCACCGACGAGTGCAGCTCGCCGCGCGCCAGAGC
CGGCTTGAGCATGTTGCCGCGTCCATGGCACCTTCCGCTTGCCGCGCGCCGACCATGGTGTGC
AGTTTCGTGATGAACAGGATGACCCGGCCTTCTGCTTGGCCAGTTTCGTTGAGGACCGCCTTCA
GGCGTTCTCGAACTCGCCGCGAACTTGGCACCGGCGATCAGCGCCCCCATGTCCAGGGCCAG
CAGGCGCTTGTCTTGGAGGCGTCCGGCACTTCGCCGTTGATGATGCGCTGGGCCAGGCCCTCG
ACGATGGCGGTCTTGCCGACGCGGGGTTCCCGGATCAGCACCGGGTGTCTTGGTCCGCGGCT
GCAGGACCTGGATGGTCCGGCGGATCTCGTCGTGCGACCGATCACCGGGTCGAGCTTGCCTTC
CTCGGCGCGCTTGGTCATGTGACGGTGTACTTGTCCAGCGCCTGGCGCGACTCCTCGACGTTTC
GGTTCGTTACCCGCTTCGCCGCCACGCAGGTTGGCCACGGCATTCTCCAGCGCCTTGCGCGACA
CGCCCTGGCCGAGCAGCAGCTTGCCGAGCCTGGTGTCTCGTCCATCGCGGCCAGCAATACCAG
CTCGCTGGAGATGAACTGGTCGCCCTTCTGCTGGGCCAGGCGGTGAGCCTGGTTGAGCAGGCGT
GCGAGATCCTGGGACAGGTTACGTCGCCGCTCGGGCTCTGGATCTTCGGCAGCGCGTCGAGTT
CTTTGTTGAGGCGGCTGCGCAGGGCGGCGATATCGAAGCCGACCTGCATCAGCAGGGGCTTGAT
CGAACCGCCTTGCTGCTCGAGCAGGGCGGAAAGCAGGTGCACCGGCTCGATGGCCGGATGGTCA
TGGCCAACGGCCAGGACTGGGCGTCGAGAGCGCCAGTTGCAGCTTGCTGGTCAAACGGTCTA
TTCGATGGGTGCTCCTTCTCTATAG

FIG. 56

clpB protein SEQ ID NO: 279

MRIDRLTSKLQLALSDAQSLAVGHDHPAIEPVHLLSALLEQQGGSIKPLLMQVGFDAALRSGL
NKELDALPKIQSPTGDVNLSDQLARLLNQADRLAQQKGDQFISSELVLLAAMDENTRLGKLLLG
QGVSRKALENAVANLRGGEAVNDPNVEESRQALDKYTVDMTKRAEEGKLDPVIGRDDEIRRTIQ
VLQRRTKNNPVLIGEPGVGKTAIVEGLAQRIINGEVPDGLKDKRLLALDMGALIAGAKFRGEFE
ERLKAVLNELGKQEGRVILFIDELHTMVGAGKAEGAMDAGNMLKPALARGELHCVGATTLDEYR
QYIEKDAALERRFQKVLVDEPSEEDTIAILRGLKERYEVHHGVSITDGAIIAAAKLSHRYITDR
QLPDKAIDLIDEAASRIRMEIDSKPEELDRRLIQLKIEREALKKEDDEATRKRRLAKLEEDI
VKLEREYADLEEIWKSEKAQVQSAQIQKIEQAKQEMEAARRKGDLESMARIQYQTIPDLERS
LQMVDQHGKTENQLLRNKVTDDEEIAEVVSKWTGIPVSKMLEGEREKLLRMEQELHRRVIGQDEA
VVAVSNAVRRSRAGLADPNRPSGSFLFLGPTGVGKTELCKALAEFLFDTEEALVRIDMSEFMK
HSVARLIGAPPGYVGFEEGGYLTEAIRRKPYSVLLDEVEKAHPDVFNILLQVLEDGRLTDSHG
RTVDFRNTVVMTSNLGSAQIQELAGDREAQRAAVMDAVNAHFRPEFINRIDEVVVFEPLAREQ
IAGIAEIQLGRLRKRLAERELSLELSQEALDKLIAVGFDVPYGARPLKRAIQRWIENPLAQLIL
AGKFAPGASISAKVEGDEIVFA.

FIG. 57

clpB DNA SEQ ID NO: 120

ATGCGAATAGACCGTTTGGACCAGCAAGCTGCAACTGGCGCTCTCCGACGCCCAGTCCCTGGCC
GTTGGCCATGACCATCCGGCCATCGAGCCGGTGCACCTGCTTTCCGCCCTGCTCGAGCAGCAA
GGCGGTTTCGATCAAGCCCCTGCTGATGCAGGTGGCTTCGATATCGCCGCCCTGCGCAGCGGC
CTCAACAAAGAACTCGACGCGCTGCCGAAGATCCAGAGCCCGACCGGCGACGTGAACCTGTCC
CAGGATCTCGCACGCCTGCTCAACCAGGCTGACCGCCTGGCCCAGCAGAAGGGCGACCAAGTTC
ATCTCCAGCGAGCTGGTATTGCTGGCCGCGATGGACGAGAACACCAGGCTCGGCAAGCTGCTG
CTCGGCCAGGGCGTGTCGCGCAAGGCGCTGGAGAATGCCGTGGCCAACCTGCGTGGCGGCGAA
GCGGTGAACGACCCGAACGTGAGGAGTCGCGCCAGGCGCTGGACAAGTACACCGTCGACATG
ACCAAGCGCGCCGAGGAAGGCAAGCTGACCCGGTGATCGGTGCGGACGACGAGATCCGCCGG
ACCATCCAGGTCTGTCAGCGGCGGACCAAGAACAACCCGGTGCTGATCGGCGAACCCGGCGTC
GGCAAGACCGCCATCGTCGAGGGCCTGGCCCAGCGCATCATCAACGGCGAAGTGCCGGACGGC
CTCAAGGACAAGCGCCTGCTGGCCCTGGACATGGGGGCGCTGATCGCCGGTGCCAAGTTCGCG
GGCGAGTTCGAGGAACGCCTGAAGGCGTCTCAACGAACTGGGCAAGCAGGAAGGCCGGGTC
ATCCTGTTTCATCGACGAACTGCACACCATGGTCGGCGCCGGCAAGGCGGAAGGTGCCATGGAC
GCCGGCAACATGCTCAAGCCGGCTCTGGCGCGCGGCGAGCTGCACTGCGTCGGTGCTACTACC
CTCGACGAGTATCGCCAGTACATCGAGAAGGATGCCGCGCTGGAGCGCCGCTTCCAGAAGGTG
CTGGTGGACGAACCGAGCGAGGAAGACACCATCGCCATCCTCCGTGGCCTCAAGGAACGCTAT
GAAGTGCAACACGGGGTGAGCATCACCGACGGCGCGATCATCGCCGCGGCCAAGCTGTCGCAC
CGTACATCACCGATCGGCAACTGCCGGACAAGGCCATCGACCTGATCGACGAGGCCGCCAGC
CGCATCCGCATGGAGATCGACTCCAAGCCGGAGGAAGTGGATCGTCTCGACCGTCGCTGATC
CAGCTGAAGATCGAGCGCGAGGCGCTGAAGAAGGAAGACGACGAAGCCACCAGGAAGCGCCTG
GCCAAGCTGGAGGAGGATATCGTCAAGCTCGAGCGCGAATACGCCGACCTCGAGGAGATCTGG
AAGTCCGAGAAGGCCGAGGTGCAGGGCTCGGCGCAGATCCAGCAGAAGATCGAGCAGGCCAAG
CAGGAGATGGAGGCGGCGCGGCGCAAGGGCGACCTCGAGAGCATGGCGCGCATCCAGTACCAG
ACCATCCCGGACCTGGAACGCGAGCCTGCAGATGGTCGACCAGCACGGCAAGACCGAGAACCAG
TTGCTGCGCAACAAGGTGACCGACGAGGAAATCGCCGAAGTGGTTTCCAAGTGGACCGGTATC
CCGGTGTGCAAGATGCTCGAGGGCGAGCGCGAGAAGCTGCTGCGCATGGAGCAGGAGCTGCAT
CGGCGAGTGATCGGCCAGGACGAGGCGGTAGTCGCCGTGTCCAACGCCGTGCGCCGTTTCGCGC
GCCGGCCTCGCCGATCCGAACCGGCCGAGCGGCTCGTTCTTCTTCTCGGCCCGACCGGGGTG
GGCAAGACCGAGTTGTGCAAGGCGCTGGCCGAGTTCTCTTTCGATACCGAGGAGGCGCTGGTG
CGGATAGATATGTCCGAGTTTCATGGAGAAACACTCGGTGGCCCGCCTGATCGGCGCGCCTCCG
GGTACGTGGCTTCGAGGAAGGCGGCTACCTGACCGAGGCGATCCGCCGCAAGCCCTACTCG
GTGGTGCTGCTGGACGAGGTGGAGAAGGCCCATCCGGATGTATTCAACATTCTCCTCCAGGTG
CTCGAGGACGGACGCCTGACCGACAGTCACGGGCGTACGGTGGACTTCCGCAACACCGTGGTG
GTGATGACCTCCAACCTCGGTTTCGGCGCAGATCCAGGAGCTGGCCGGCGACCGCGAGGCGCAA
CGTGCCGCGAGTGATGGACGCGGTCAATGCGCACTTCCGTCCGGAATTCATCAACCGGATCGAC
GAAGTGGTGGTGTTCGAGCCGCTGGCTCGCGAGCAGATCGCCGGCATCGCCGAGATCCAGCTC
GGTCGCCTGCGCAAGCGCCTGGCCGAGCGCGAGCTGAGCCTGGAACCTGAGCCAGGAGGCGCTG
GACAAGCTGATTGCCGTGGCTTCGACCCGGTCTATGGCGCACGCCCGCTGAAGCGGGCCATC
CAGCGCTGGATCGAGAACCCTGCTGGCGCAACTGATCCTGGCCGGCAAATTCGCGCCGGGTGCC
AGTATCTCGGCGAAGGTGGAAGGCGACGAGATCGTCTTCGCCTGA

FIG. 58

ORF ID	Strand	Left end	Right end	ORF length (aa)	G+C content (%)	Location prediction	Gene name	Gene function	Protein with the highest identity (Gene Name / Species Strain)	E-value (% identity)	GenBank accession no.	Proteins with lesser identity Cut off 30%
	+	801	876		51				IRNA-lys / <i>P. aeruginosa</i> PAO1		AE004531	
RS01	-	959	1,280	pseudogene	57			Hypothetical protein	PA0977 / <i>P. aeruginosa</i> PAO1	8E-37 (86)	AAG04366	
RS02	-	1,247	2,527	426	60	cytoplasm	xerC	Integrase	XerC / <i>P. aeruginosa</i> SG17M	0.0 (94)	AAG02084	STY4666
RS03	-	2,524	2,901	125	56	cytoplasm		Hypothetical protein	XF1753 / <i>X. fastidiosa</i> 9a5c	3E-34 (49)	AAF84562	XAC2196, XCC3121, STY4665
RS04	-	2,988	3,830	280	62	cytoplasm		Putative transposase	PA0978 / IS222 / <i>P. aeruginosa</i> PAO1	1E-154 (96)	AAG04367	PA1938, RSC2313
RS05	-	3,830	4,138	103	60	cytoplasm		Putative transposase	PA0979 / IS222 / <i>P. aeruginosa</i> PAO1	6E-50 (99)	AAG04368	PA1937, RSC2314
RS06	+	4,370	4,654	94	45	inner or outer membranes		Hypothetical protein	PA0980 / <i>P. aeruginosa</i> PAO1	4E-50 (97)	AAG04369	
RS07	+	4,691	5,314	207	46	outer membrane and periplasm		Hypothetical protein	PA0981 / <i>P. aeruginosa</i> PAO1	E-114 (99)	AAG04370	
RS08	-	5,420	5,737	105	50	outer membrane and periplasm		Hypothetical protein	PA0673 / <i>P. aeruginosa</i> PAO1	2E-05 (37)	AAG04062	
RS09	+	5,849	6,052	67	57	cytoplasm			No significant similarity			
RS10	+	6,108	6,326	72	58	cytoplasm			No significant similarity			
RS11	+	6,540	6,989	149	53	inner membrane		Acetyltransferase	PP0651 / <i>P. putida</i> KT2440	2E-75 (95)	AAN66276	
RS12	-	7,118	8,218	366	59	outer membrane and periplasm		Transposase	PP3964 / ISPpu14 ORF3 / <i>P. putida</i> KT2440	0.0 (98)	AAN69558	
RS13	+	8,253	8,822	189	61	cytoplasm		Putative transposase	PA0987 / <i>P. aeruginosa</i> PAO1	4E-88 (85)	AAG04376	XAC2424
RS14	+	8,959	11,022	687	59	cytoplasm	exoU (pepA)	cytotoxin (type III secretion system effector)	ExoU / <i>P. aeruginosa</i> PA103	0.0 (100)	AAC38269	
RS15	+	11,019	11,432	137	56	inner membrane	spcU	ExoU chaperone	SpcU / <i>P. aeruginosa</i> PA103	4E-62 (94)	AAC16024	

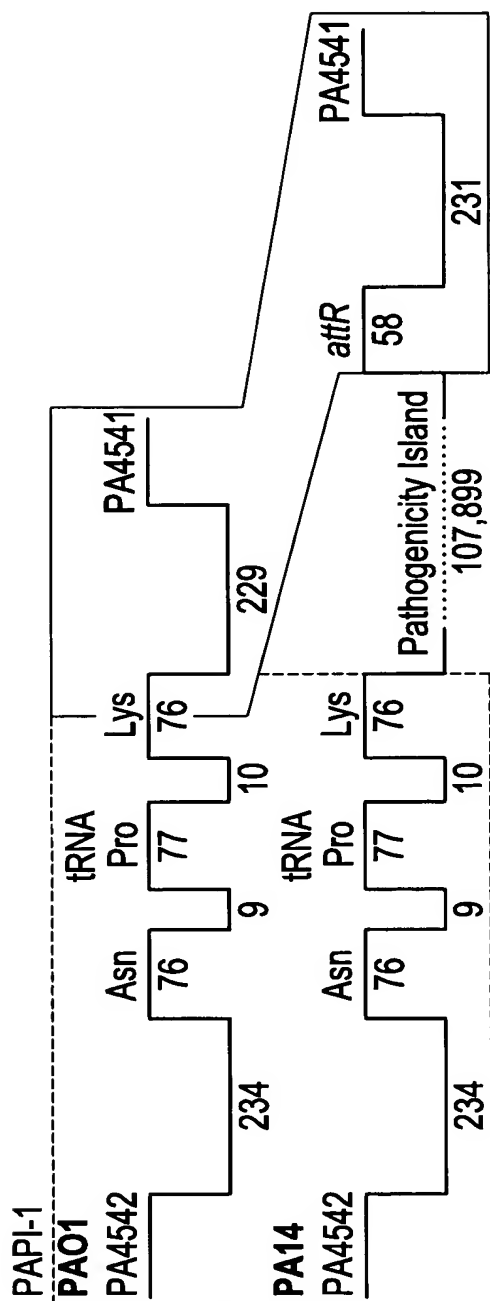
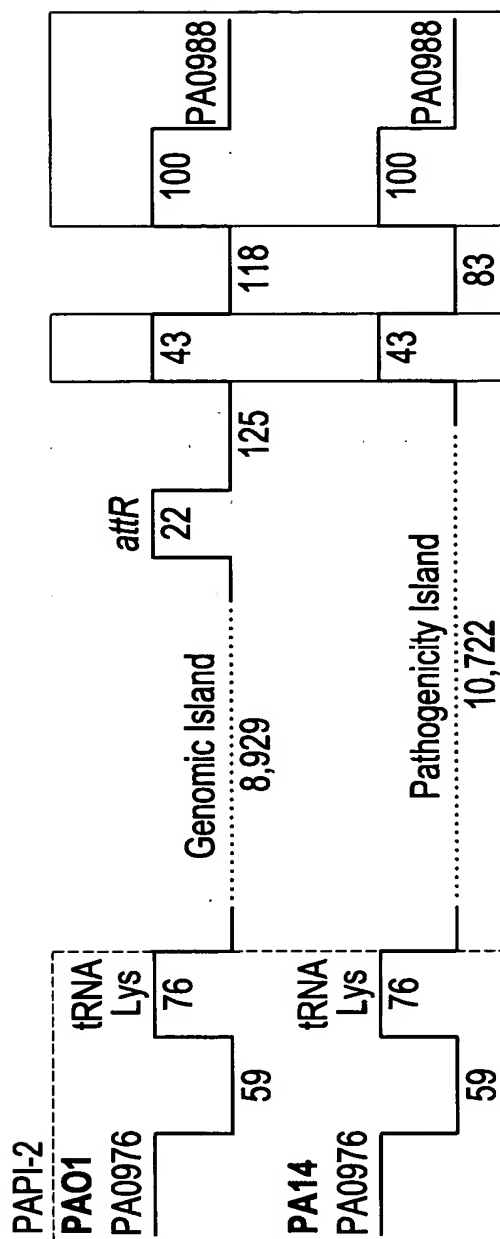


FIG. 59B



107/118

FIG. 60A

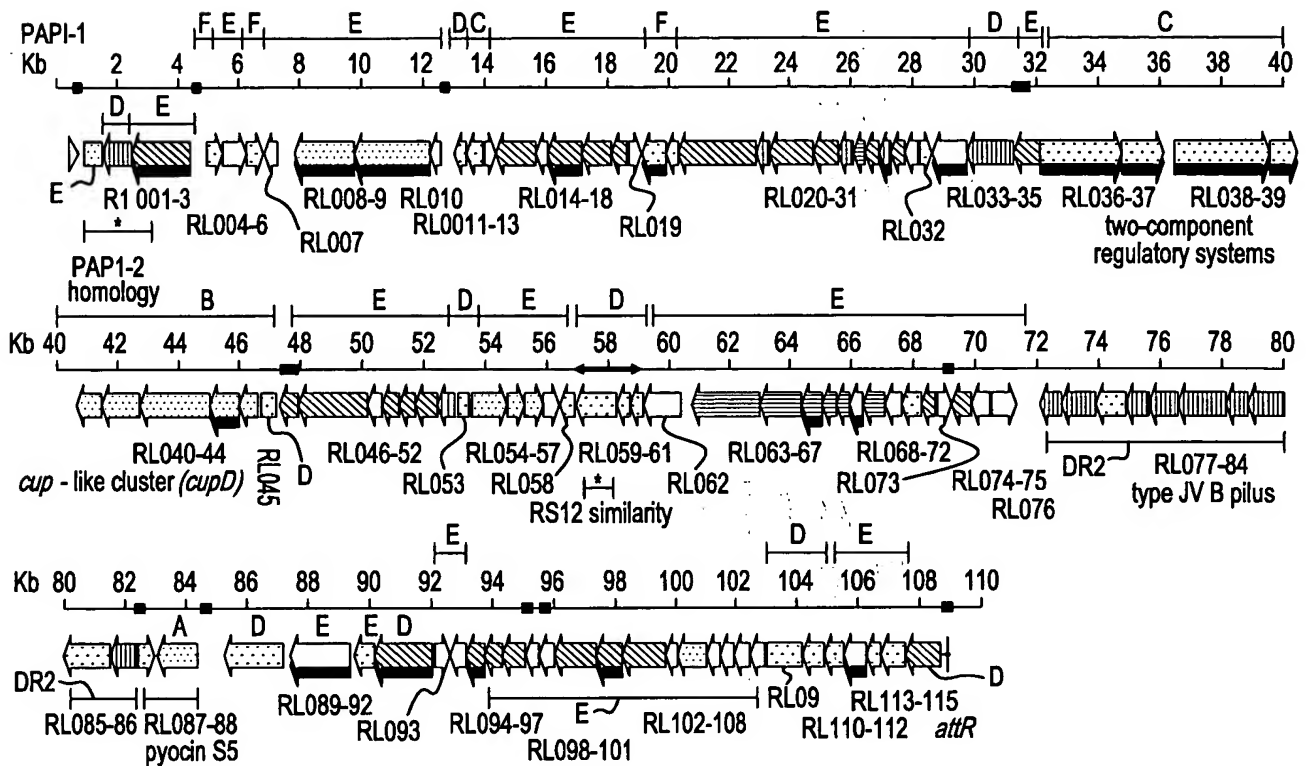


FIG. 60B

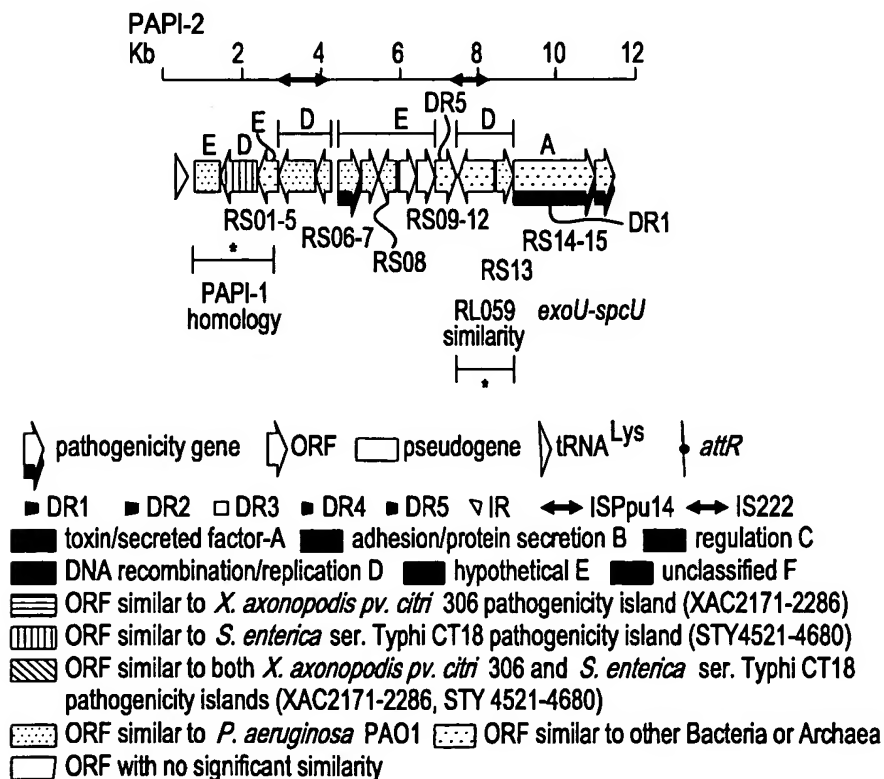


FIG. 61

	▼ DR1			▼ DR2			▼ DR2			DR3 ▼			▼ DR3			DR1 ▼	
kb	0	10	20	30	40	50	60	70	80	90	100						
PA14	++	+	+	+	+		+	+	+		++						
CF2	++	+	+	+	+		+	+	+		++						
CF6	++	+	+	+	+		+	+	+		++						
PA037	N+	+	+	+	+		N	+	+		+N						
CF26	++	+	+				+	+	+		++						
CF29	++	+	+	+	+		+										
PAK	++																
PAO1									+								

109/118

FIG. 62A

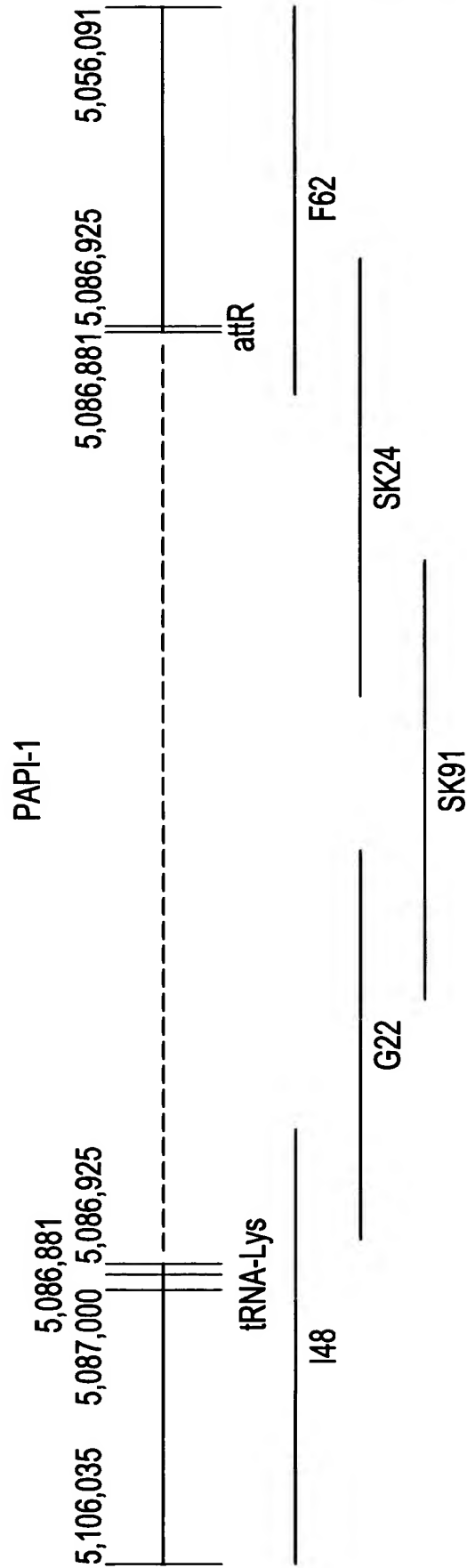


FIG. 62B

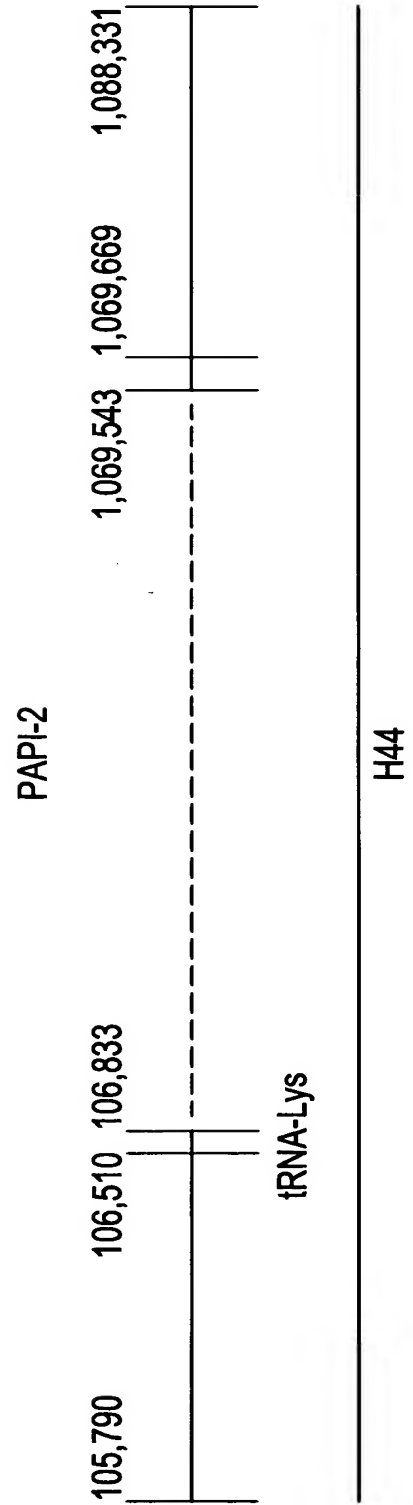


FIG. 63

Strain name*	% Mouse mortality†	Growth in <i>Arabidopsis</i> leaf‡	Closest published homologue (organism / GenBank accession no.)
PA14	100	4.9×10^6	
RL003§	41	2.3×10^5	<i>P. syringae</i> pv. <i>tomato</i> DC3000 / AAO54371
RL008	38	4.1×10^6	<i>M. acetivorans</i> C2A / AAM05538 and <i>P. aeruginosa</i> PAO1 / AAG05323
RL009	31	1.3×10^4	<i>P. aeruginosa</i> PAO1 / AAG05327
RL016	100	2.8×10^4	<i>P. syringae</i> pv. <i>tomato</i> DC3000 / AAO54383
RL020	50	3.4×10^5	protein-disulfide isomerase, <i>P. aeruginosa</i> PAO1 / AAG04371
RL022	88	3.3×10^6	<i>P. syringae</i> pv. <i>tomato</i> DC3000 / AAO54394
RL029	38	9.4×10^4	<i>P. aeruginosa</i> C / AAN62148
RL033	25	4.9×10^4	no significant similarity
RL036	44	1.9×10^5	two-component sensor <i>P. aeruginosa</i> PA14 / AAM15532
RL037	43	1.2×10^5	two-component regulator <i>pvrR</i> , <i>P. aeruginosa</i> PA14 AAM15533
RL038	31	4.4×10^4	two-component sensor <i>rscC</i> , <i>S. typhimurium</i> LT2 / AAL21172
RL039	31	2.7×10^5	two-component regulator <i>rscB</i> , <i>E. coli</i> O157:H7 EDL933 / AAG57352
RL043	75	1.7×10^6	probable pili assembly chaperone <i>cupA2</i> , <i>P. aeruginosa</i> PAO1 / AAG05517
RL054	63	NT¶	<i>P. aeruginosa</i> PAO1 / AAG05610
RL062	78	NT¶	no significant similarity
RL065	63	4.5×10^5	<i>X. axonopodis</i> pv. <i>citri</i> 306 / AAM37094
RL068	56	2.6×10^5	no significant similarity
RL090	67	2.7×10^4	no significant similarity
RL092	0	1.3×10^5	topoisomerase I TopA, <i>X. fastidiosa</i> 9a5c (plasmid pXF51) / AAF85572
RL095	50	5.3×10^5	single-stranded DNA binding protein Ssb, <i>P. aeruginosa</i> C / AAN62318
RL101	38	1.8×10^6	<i>Pseudomonas</i> sp. B13 / CAD60668
RL112	38	1.6×10^4	no significant similarity
RS06	100	1.8×10^5	<i>P. aeruginosa</i> PAO1 / AAG04369

FIG. 64

	Positions	Length (bp)	Number of identical bp	Genes between the repeats
DR1	744-805 108,700-108,762	63	59	PAP1-1 (108 Kb)
DR2	31,587-32,248 47,100-47,761	662	654	two component regulatory systems and <i>cup</i> -like cluster (<i>cupD</i>)
DR3	82,574-82,821 85,296-85,540	248	231	pyocin S5 and associated immunity protein
DR4	95,301-95,357 95,358-95,414	57	50	none
DR5	95,767-95,824 95,825-95,881	58	54	none
IR	4,527-4,594 12,825-12,892	68	56	pathogenicity genes and Archaea homologous genes

FIG. 65

IS name	PAPI-1		PAPI-2		Characteristics of IS	
	Position	Length (bp)	Position	Length (bp)	Original length (bp)	IS family
ISPpu14	56,778- 59,119	2,341	7,034- 7,999	966	2,383	IS66
IS222	-	-	2,980- 4,201	1,222	1,232	IS3

FIG. 66

Function	Prototype name	Type IVB (PAPI-1 in PA14)	Type IVA (PAO1)	Xcp (PAO1)	Hxc (PAO1)	Hpl (PAO1)	Other homologues in PAO1 genome
ATPase	<i>puE</i>	RL082 (<i>pilQ2</i>)	<i>pilB</i> <i>pilT</i> <i>pilU</i>	<i>xcpR</i>	<i>hxcR</i>	<i>hplR</i>	<i>hvbA</i> , <i>hxrA</i>
Peptidase	<i>puO</i>	RL079 (<i>pilT2</i>)?	<i>pilD</i>	<i>pilD</i>	<i>pilD</i>	<i>pilD</i> ?	
Major pilin	<i>puG</i>	RL080 (<i>pilS2</i>)	<i>pilA</i>	<i>xcpT</i>	<i>hxcT</i>	<i>hplT</i>	
Minor pilin	<i>puH</i> <i>puI</i> <i>puJ</i> <i>puK</i>	RL077 (<i>pilM2</i>) RL078 (<i>pilV2</i>) RL083 (<i>pilP2</i>) RL086 (<i>pilL2</i>)	<i>pilE</i> <i>fimU</i> <i>fimT</i>	<i>xcpU</i> <i>xcpV</i> <i>xcpW</i> <i>xcpX</i>	<i>hxcU</i> <i>hxcV</i> <i>hxcW</i> <i>hxcX</i>	<i>hplU</i> <i>hplV</i> <i>hplW</i> <i>hplX</i>	
Inner membrane protein	<i>puIF</i> <i>puIC</i> <i>puIL</i> <i>puIM</i>	RL081 (<i>pilR2</i>) RL084 (<i>pilO2</i>)	<i>pilC</i>	<i>xcpS</i> <i>xcpP</i> <i>xcpY</i> <i>xcpZ</i>	<i>hxcS</i> <i>hxcP</i> <i>hxcY</i> <i>hxcZ</i>	<i>hplS</i>	<i>xqhA</i>
Secretin	<i>puD</i>	RL085 (<i>pilN2</i>)	<i>pilQ</i>	<i>xcpQ</i>	<i>hxcQ</i>		<i>xqhA</i> , <i>xqhB</i> , <i>xqhC</i>

114/118
FIG. 67

ORF 7 (SEQ ID NO: 280)

LEFGSATWTRTRDPMINSHLLYRLSYRGTSFFQPWTLVLLDSRLRGAPFYGCARACQPSDPKSFSSFSTSDKTAL
PLHAAALSRLPDAHEKAPPKRGFPKPPKRSGEDDLVAFHLRRDTGTRREFAGQDQLRQRLVDPALDGPLQRACAI
DRVEADGNQLVQRLLAQFQAQLALGQALAQATELDLGDAGDLLASQRLEHHHFVDPVDEFRTVEVRIDRVHHCGLTR
LAVAGQLDLRRTEVGGHHHGVAEVHRTPTVTVGQASVLEHLEENVEYIRMGLLHLVQQHHRVGLAADRLGQVAAF
LEADVARRRADQAGHRVFLHELGHYYPHQRLLGIEEELGQRLAQLGLAHPGRAEEERAARPVRIGEAGARTAHGV
GHGDYRLVLADHSPMQLLLHAQQLLALALEHLRHRDTGPLGNHFGDFLVGHLVAQQLVLGLAVLVDHLQAQFQVRD
GLVLDARHALEVALAPRRLHLLGLLDLLDLRRALHLGLLGLPDLLEVGVFALELDDILLQLGQALPGGFVVFL
QRLALDLQLDQATVETIQFLRLGVDLHADAAGGLVDQVDGLVRQLPIGDVAVRQLGRGDDRAVGDAHPVVHFI AFL
EATEDGDGVFLARFVHQHLLLEAALQIRGILLDVLAILEVGSSTDAVQLAARQSRLEHVAGVHGTFRLAGADHGVQFV
DEQDDPAFLLAQFVEDRLQAFLELAAELGTGDQRPHVQGGQALVLEAVRHFAVDDALGQALDDGGLADAGFADQHR
VVLGPPLQDLGDPADLVVATDHRVELAFLGALGHVDGVLVQRLARLLDVRVVRHFAATQVGHGILQRLARHALAEQ
QLAEPGVLVHRGQQYQLAGDELVALLLGQAVSLVEQACEILGQVHVAGRALDLRQRVEFFVEAAAQGGDIEADLHQ
QGLDRTALLLEQGGKQVHRLDGRVMANGQGLGVGERQLQLAGQTVYSHGSSFL

ORF7 (SEQ ID NO: 281)

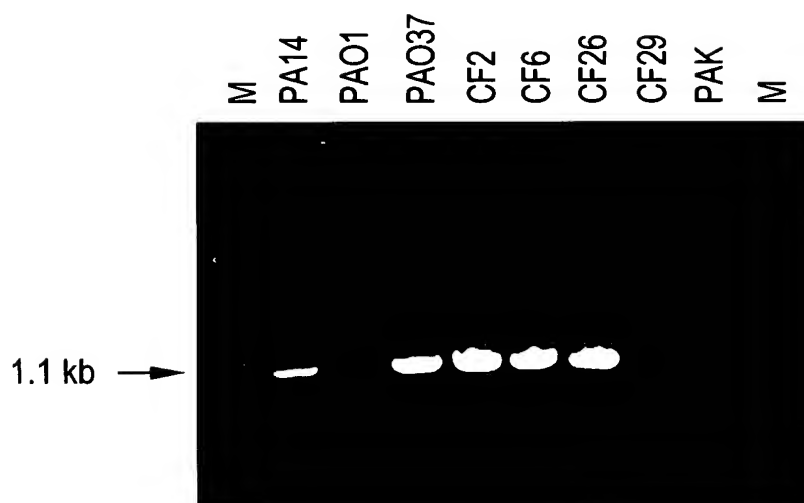
TTGGAATTTGGCTCCGCGACCTGGACTCGAACCAGGGACCCAATGATTAACAGTCATTTGCTCTACCGACTGAGCT
ATCGCGGAACGTCTTTCTTCCAACCTGGACGCTTCCGGTGTGCTGGATTTCGCGTCTCAGAGGCGCGCCATTTTA
CGGATGCGCGCGGGCATGTCAACCTCTGATCCAAAAGTTTTTCTTCTTTTCCACGAGCGACAAAACGGCCCTT
CCACTGCATGCGGCAGCGCTCTCGCGCTACCGGACGCCCATGAAAAGCCCCGCGAAGCGGGGCTTTCCCTGTC
CGCCCCCGAAGAGGTGAGGCGAAGACGATCTCGTCGCTTCCACCTTCGCCGAGATACTGGCACCCGGCGCGAATT
TGCCCGCCAGGATCAGTTGCGCCAGCGGTTCTCGATCCAGCGCTGGATGGCCCGCTTCAGCGGGCGTGCGCCATA
GACCGGGTCGAAGCCGACGGCAATCAGCTTGTCAGCGCCTCCTGGCTCAGTTCCAGCTCAGCTCGCGCTCGGCC
AGGCGCTTGCGCAGGCGACCGAGCTGGATCTCGGCGATGCGCGCGATCTGCTCGCGAGCCAGCGGCTCGAACACCA
CCACTTCGTCGATCCGGTTGATGAATCCGGACGGAAGTGCATGACCGCGTCCATCACTGCGGCACGTTGCGC
CTCGCGGTGCGCCGCGCAGCTCCTGGATCTGCGCCGAACCGAGGTTGGAGGTATCACCACCACGGTGTGCGGAAG
TCCACCGTACGCCCCGTGACTGTGCGTCAAGCGTCCGTCTCGAGCACCTGGAGGAGAATGTTGAATACATCCGGAT
GGGCTTCTCCACCTCGTCCAGCAGCACACCGAGTAGGGCTTGCGCGGATCGCCTCGGTCAAGTAGCCGCTTC
CTCGAAGCCGAGCTAGCCCGGAGGCGCGCCGATCAGGCGGGCCACCGAGTGTTTCTCCATCACTCGGACATATCT
ATCCGCACACGCGCTCCTCGGTATCGAAGAGGAACCTCGGCCAGCGCTTGACACAACCTCGGTCTTGCCACCCCGG
TCGGGCGGAGGAAGAGGAACGAGCGCTCGGCCGTTGCGATCGGCGAGGCGCGCGCGAAGCGGCGACGGCGTT
GGACACGGGCGACTACCGCTCGTCTGCGCGATCACTCGCCGATGCAGCTCCTGCTCCATGCGCAGCAGCTTCTCG
CGCTCGCCCTCGAGCATCTTCGACACCGGGATACCGGTCCACTTGGAACCACTTCGGCGATTTCTCTCGTGGTCA
CCTTGTTGCGCAGCAACTGGTTCTCGGTCTTGCGTGTGCTGACCATCTGCAGGCTGCGTTCAGGTCCGGGAT
GGTCTGGTACTGGATGCGCGCCATGCTCTCGAGGTGCGCTTGCGCCGCGCGCTCCATCTCCTGCTTGGCCTGC
TCGATCTTCTGCTGGATCTGCGCCGAGCCCTGCACCTCGGCCTTCTCGGACTTCCAGATCTCCTCGAGGTGCGCGT
ATTCGCGCTCGAGCTTGACGATATCCTCCTCAGCTTGCCAGGCGCTTCTGGTGGCTTCGTCGCTTCTCTTCTT
CAGCGCTCGCGCTCGATCTTCAGCTGGATCAGGCGACGGTCGAGACGATCCAGTTCTCCTCGGCTTGAGTCGATC
TCCATGCGGATGCGGCTGGCGGCCCTCGTCGATCAGGTGCGATGGCCTTGTCGGGAGTTGCCGATCGGTGATGTAGC
GGTGCGACAGCTTGCCCGCGCGGATGATCGCGCGCTCGGTGATGCTACCCCGTGGTGCACTTCATAGCGTTCTCT
GAGGCCACCGAGGATGGCGATGGTGTCTTCTCGCTCGGTTCGCTCCACCGACACCTTCGGAAGCGCGCTCCAGC
GCGGCATCCTTCTCGATGTACTGGCGATACTCGTCGAGGGTAGTAGACCGACGCGAGTCGAGCTCGCCGCGCGCCA
GAGCCGGCTTGAGCATGTTGCCGGCTCCATGGCACCTTCGCGCTTGCCGGCGCCGACCATGGTGTGAGTTTCGTC
GATGAACAGGATGACCGGCCCTTCTGCTTGCCAGTTGCTTGAGGACCGCTTCAGGCGTTCCTCGAAGTCCGCG
CGGAACCTTGGCACCGGCGATCAGCGCCCCCATGTCCAGGGCCAGCAGGCGCTTGCTCTTGAGGCGCTCCGGCACTT
CGCGCTTGATGATGCGCTGGGCCAGGCCCTCGACGATGCGCGCTTGCCGACGCGGGTTCGCGCATCAGCACCGG
GTGTTCTTGTTGCTCGCGCGCTGACGAGACCTGGATGGTCCGCGGATCTCGTCTGCGCGACCGATCACCGGCTCAGC
TTGCTTCTCGCGCGCTTGCTCATGTGCGACGGTGTACTTGTCAGCGCTTGCGCGGACTCCTCGACGTTCCGGT
CGTTCACCGCTTCGCGGCCACGAGGTTGGCCACGGCATCTCCAGCGCTTGCGCGACACGCCCTGGCCGAGCAG
CAGCTTGCCGAGCCTGGTGTCTCGTCCATCGCGCCAGCAATACCAGCTCGTGGAGATGAAGTGGTCGCCCTTC
TGCTGGGCGAGGCGGTGAGCTGTTGAGCAGGCGTGCAGATCCTGGACAGGTTACGTCGCGCGTGGGCTCT
GGATCTTCCGCGAGCGCTGAGTTCTTTGTTGAGGCGCTGCGCAGGGCGGATCGAAGCCGACCTGCATCAG
CAGGGCTTGATCGAAGCGCTTGCTGCTCGAGCAGGCGGAAAGCAGGTGCACCGGCTCGATGGCGGATGGTCA
TGGCCAACGGCCAGGACTGGCGCTCGGAGAGCGCCAGTTGCAGCTTGCTGGTCAAACGGTCTATTTCGATGGGT
GTCCTTCTCTATAG

FIG. 68

(SEQ ID NO: 121)

ACGTCGGGGGCGCATTGctACGCCTGcAgAATGGTTTCAGGGCCTTAGAAACAGAAAAGCCCA
CC**TaGAC**AGGCGGGCTATTCCATATTGA**cATcAc**GTCAATGCGGGCCTAATGTTTCGGCCCAGA
CGGCTGCTAGACAAGAACCGGCGTAACACCCCTTCCTAGCCTATGCAACTCGCCCCG**T**AGAAA
ATGGTGGGTCGTGTAGGATTCGAACCTACGACCAATTGGTTAAAAGCCAAC**TGCTCTACCGAC**
TGAGCTAACGACCCAAGTATGAGGTGGTCGGGGTAGAGAGATT**CGAACTCCCGACATCCTGCT**
CCCAAAGCAGGCGCGCTACCGGACTGCGCTATACCCCGATTGGAATTTGGCTCCGCGACCTGG
ACTCGAACCAGGGACCCAATG

FIG. 69



117/118 FIG. 70A

CLUSTAL W (1.82) multiple sequence alignment

```

CF2 (SEQ ID NO: 122) -----GATGAAGG-ACCCGAGCGGAACATCCATCTCAC 32
PAO37 (SEQ ID NO: 123) -----GATGAAGG-ACCCGAGCGGAACATCCATCTCAC 32
CF6 (SEQ ID NO: 124) -----GATGAAGGCACCCGAGCGGAACATCCACCTCAC 33
PA14 (SEQ ID NO: 125) TTTCCGGCACACCTGGCCACCGACTTGATGAAGGCACCCGAGCGGAACATTACCTCAC 60
CF26 (SEQ ID NO: 126) -----GATGAAGG-ACCCGAGCGGAACATCCACCTCAC 32
                        *****

CF2 GAAGTGCCTGCTCAACCACTCGAATATCCAGACCACCATGAGCTACATCGAGGCCGACTA 92
PAO37 GAAGTGCCTGCTCAACCACTCGAATATCCAGACCACCATGAGCTACATCGAGGCCGACTA 92
CF6 GAAGTGCCTGCTCAACCACTCGAATATCCAGACCACCATGAGCTACATCGAGGCCGACTA 93
PA14 GAAGTGCCTGCTCAACCACTCGAATATCCAGACCACCATGAGCTACATCGAGGCCGACTA 120
CF26 GAAGTGCCTGCTCAACCACTCGAATATCCAGACCACCATGAGCTACATCGAGGCCGACTA 92
                        *****

CF2 CGACCACATGCGTGCCGTGCTGCATGCCAGAAGCCTGGCCCAAGGAGCGCTGGAGAACGT 152
PAO37 CGACCACATGCGTGCCGTGCTGCATGCCAGAAGCCTGGCCCAAGGAGCGCTGGAGAACGT 152
CF6 CGACCACATGCGTGCCGTGCTGCATGCCAGAAGCCTGGCCCAAGGAGCGCTGGAGAACGT 153
PA14 CGATCACATGCGTGCCGTGCTGCATGCTAGAAGCCTGGCCCAAGGCGCGCTGGAGAATGT 180
CF26 CGATCACATGCGTGCCGTGCTGCATGCTAGAAGCCTGGCCCAAGGCGCGCTGGAGAATGT 152
                        *** *****

CF2 CAGGAAGGTGGATTACAGCGGCTCCCCGCAAGCCTCTGCCAAACCGAAGCCATGCGGGCA 212
PAO37 CAGGAAGGTGGATTACAGCGGCTCCCCGCAAGCCTCTGCCAAACCGAAGCCATGCGGGCA 212
CF6 CAGGAAGGTGGATTACAGCGGCTCCCCGCAAGCCTCTGCCAAACCGAAGCCATGCGGGCA 213
PA14 CAGGAAGGTGGATTACAGCGGCTCCCCGCAAGCCTCTGCCAAACCGAAGCCATGCGGGCA 240
CF26 CAGGAAGGTGGATTACAGCGGCTCCCCGCAAGCCTCTGCCAAACCGAAGCCATGCGGGCA 212
                        *****

CF2 ACCTCTCGCTCGAATGGGTGAAGTACCGCCGCCGAGGCCAGGACAGAACCTGCAGAAC 272
PAO37 ACCTCTCGCTCGAATGGGTGAAGTACCGCCGCCGAGGCCAGGACAGAACCTGCAGAAC 272
CF6 ACCTCTCGCTCGAATGGGTGAAGTACCGCCGCCGAGGCCAGGACAGAACCTGCAGAAC 273
PA14 ACCTCTCGCTCGAGTGAAGTGAAGCGCCGCCACCGGAGGCCAGGACAGACCTGCAGAAC 300
CF26 ACCTCTCGCTCGAGTGAAGTGAAGCGCCGCCACCGGAAGCCAGGACAGACCTGCAGAAC 272
                        *****

CF2 AAGGGAGCACATACCAGGGACAGGCATTACAGGGAGGTCCAACCGTGCGGGAAGAAG---C 329
PAO37 AAGGGAGCACATACCAGGGACAGGCATTACAGGGAGGTCCAACCGTGCGGGAAGAAG---C 329
CF6 AAGGGAGCACATACCAGGGACAGGCATTACAGGGAGGTCCAACCGTGCGGGAAGAAG---C 330
PA14 AAGGGAGCACACGCCAGGGACAGGCATTACAGGGAGGTCCAACCGTGCGGGAAGCAGATGC 360
CF26 AAGGGAGCACACGCCAGGGACAGGCATTACAGGGAGGTCCAACCGAGTGGAAGCAGAAAGC 332
                        *****

CF2 GCTACCACAGCCACCTGACACCTTCGACCAAAGCGTGCTGTTCACTCTGATGGCTCAACA 389
PAO37 GCTACCACAGCCACCTGACACCTTCGACCAAAGCGTGCTGTTCACTCTGATGGCTCAACA 389
CF6 GCTACCACAGCCACCTGACACCTTCGATCAAAGCGTGCTGTTCACTCTGATGGCTCAACA 390
PA14 GCTACCACAGCCACCTGACACCTTCGAAACCAAGCGTGCTGTTCACTCTGATGGCTCAAAA 420
CF26 GCTACCACAGCCACCTGACACCTTCGAGCAAAGCGTGCTGTTCACTCTGATGGCTCAACA 392
                        *****

CF2 CTTATCGAACCGTGCCGCTCGGCATCCGCGGCTCCCGCTGCAACAAGCGGATCTGGTGG 449
PAO37 CTTATCGAACCGTGCCGCTCGGCATCCGCGGCTCCCGCTGCAACAAGCGGATCTGGTGG 449
CF6 CTTATCGAACCGTGCCGCTCGGCATCCGCGGCTCCCGCTGCAACAAGCGGATCTGGTGG 450
PA14 CTTATCGAACCGTGCCGCTCGGCATCCGCGGCTCCCGCTGCAACAAGCGGATCAGGCGG 480
CF26 CTTATCGAACCGTGCCGCCACGACATCTGCGGCTCCCGCGCAACCAGCGGATCTGTAG 452
                        *****

CF2 ATGGGGATCTACTGCCCGAAGCAGTCTCGCCTAGCGATACCGGATCTGAAGGGCCGGCTA 509
PAO37 ATGGGGATCTACTGCCCGAAGCAGTCTCGCCTAGCGATACCGGATCTGAAGGGCCGGCTA 509
CF6 ATGGGGATCTACCGCCCGAAGCAGTCTCGCCTAGCGATACCGGATCTGAAGGGCCGGCTA 510
PA14 ATGGGGATCTGCCGCCCGAAGCAATCTCGCCTAGCGATACCGGATCTGAAGGGCCGGCTA 539
CF26 ATGGGGATCTGCCGCCCGAAGCAGCCTCGCCTAGCGATACCGGATCTGAGGGCCGGCTA 512
                        *****

```

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND
PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September 12, 2003 Serial No.: Not Yet Assigned

Page 117 of 118 Customer No.: 21559

FIG. 70B

CF2	CCGGACGAAAGGTAGCCGCGCCTCCCAGCAGTTCGCTAGGCCTGTAAGAAAAATCTGGAA	569
PA037	CCGGACGAAAGGTAGCCGCGCCTCCCAGCAGTTCGCTAGGCCTGTAAGAAAAATCTGGAA	569
CF6	CCGGACGAAAGGTAGCCGCGCCTCCCAGCAGTTCGCTAGGCCTGTAAGAAAAATCTGGAA	570
PA14	CCGGACGAAAGGTAGCCGCGCCTCCCAGCAGTTCGCTAGGCCTGTAAGAAAAATCTGGAA	599
CF26	CCAGACGAAAGGTAGCCGCGCCTCCCAGCAGATCGTGGGCCTGTAAGAAAAATCTGGAA	572
	** *****	
CF2	TTACCGAGAGCGCCTGGATTCCAGCGCCGGCATGCTGGCAGAGCCC-CGCAGTTTCACGG	628
PA037	TTACCGAGAGCGCCTGGATTCCAGCGCCGGCATGCTGGCAGAGCCC-CGCAGTTTCACGG	628
CF6	TTACCGAGAGCGCCTGGATTCCAGCGCCGGCATGCTGGCAGAGCCC-CGCAATTTCAAGG	629
PA14	TTACCGAGAGCGCCTGGATTCCAGCGCCGGCATGCTGGCAGAGCCAGCGCAATTTCAAGG	659
CF26	TTACCGAGAGCGCCTGGATTCCAGCGCCGGCATGCTGGCAGAGCCC-CGCAATTTCAAGG	631

CF2	CCAAAACCGCAGTACCCTCTGTAATCGCTGATTACGTCGGGGGCGCATTGCTACGCCTGC	688
PA037	CCAAAACCGCAGTACCCTCTGTAATCGCTGATTACGTCGGGGGCGCATTGCTACGCCTGC	688
CF6	C-GAAACCGCAGTACCCTCTGTAATCGCTGATTACGTCGAGGGGCACATTGCTACGCCTGC	688
PA14	CCAAATACCACAGTACCCTCTGTAATCGCTGATTACGTCGGGGGCGCATTGCTACGCCTGC	719
CF26	C-AAAACCGCAGTACCCTCTGTAATCGCTGATTACGTCGGGGGCGCATTGCTACGCCTGC	690
	* * ** *****	
CF2	AGAA-TGGTTTCAGGGCCTTANAAACAGAAAAGCCCACCTTAAATAGGCGGGCTATT-CC	746
PA037	AGAAATGGTTTCAGGGCCTTAGAAACAGAAAAGCCCACCTTAAATAGGCGGGCTATT-CC	747
CF6	AGAA-TGGTTTCAGAGCCT-GAAAACAGAAAAGNCCACC-TAAATAGGCGGGCTATTTC	745
PA14	AGAA-TGGTTTCAGGGCCTTAGAAACAGAAAAGCCCACC-TAGAAAGGCGGGCTATT-CC	776
CF26	AGAA-TGGTTTCAGAGCCTTANAAACAGAAAAGCCCACC-TAGATAGGCGGGCTATT-CC	747
	**** *****	
CF2	ATATT-GACATCACG-TCAATGCGGG--CCTAATGTTC--GGCCCANACGGCTG--CTGG	798
PA037	ATATT-GACATCACG-TCAATGCGGG--CCTAATGTTC--GGCCCANACGGCTG--CTGG	799
CF6	ATATTTGACATCCCG-TCAATGCGGGGCCCTAATGGTTCGGGCCCCANACGGCTTGCTTG	804
PA14	ATATT-GACATCACG-TCAATGCGGG--CCTAATGTTC--GGCCCANACGGCTG--CTAG	828
CF26	ATATT-GACATCACGGTCAATGCGGG--GCTAATGTTC-GGGCCCANACGGNTG--CAA	800
